

412018

SEARCH REQUEST FORM

Access DB# _____

Scientific and Technical Information Center

Requester's Full Name: Natalie Davis Examiner #: 78462 Date: 7-23-01
 Art Unit: 1042 Phone Number 308-6410 Serial Number: 09/509775 CRFE
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL
8E12 CM1/9809

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search claims 1-5 + 16-17

Please search SEQ ID NO: 1+2
 and

amino acids at position 14 to 226 of SEQ ID NO: 2
 amino acids at position 1 to 226 of SEQ ID NO: 2

Sept / Structure

RECEIVED
 JUL 23 2
 (STC)

STAFF USE ONLY

Searcher: <u>D. Schreiber</u>	Type of Search	Vendors and cost where applicable
Searcher Phone #: <u>308-4292</u>	NA Sequence (#) <u>1</u>	STN _____
Searcher Location: <u>CM1 12E18</u>	AA Sequence (#) <u>3</u>	Dialog _____
Date Searcher Picked Up: <u>8/13</u>	Structure (#) _____	Questel/Orbit _____
Date Completed: <u>8/20</u>	Bibliographic _____	Dr. Link _____
Searcher Prep & Review Time: <u>8</u>	Litigation _____	Lexis/Nexis _____
Clencal Prep Time: _____	Fulltext _____	Sequence Systems <u>CompuGen</u>
Online Time: <u>8</u>	Patent Family _____	WWW/Internet _____
	Other _____	Other (specify) _____

PTO-1590 (1-2000)

BEST AVAILABLE COPY

THIS PAGE BLANK (USP)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:46:43 ; Search time 127.64 Seconds
(without alignments)
3837.067 Million cell updates/sec

```

Title:
perfect score:
Sequence: 1 tggatgaagcgtctaaccgqctg.....gaatggtgaaggttaaac 780
US-09-509-775-1
780

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
1 : /SDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2 : /SDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3 : /SDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4 : /SDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5 : /SDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6 : /SDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7 : /SDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8 : /SDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9 : /SDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10 : /SDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11 : /SDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12 : /SDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13 : /SDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14 : /SDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15 : /SDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16 : /SDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17 : /SDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18 : /SDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19 : /SDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20 : /SDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21 : /SDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22 : /SDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query %		Length	DB	ID	Description		
		Match	Length						
1	780	100.0	780	20	AA35852		DNA encoding human		
2	706	90.5	1468	18	AAT66425		Human p28 coding s		
3	678	86.9	678	18	AAT66424		Human p28 coding s		
4	593	76.0	696	20	AA35854		DNA encoding rat g		
5	584	74.9	696	20	AA35853		DNA encoding mouse		
6	93.4	12.0	1056	21	AA358272		D. immitis ankyrin		
7	93.4	12.0	1056	21	AA358273		D. immitis ankyrin		
8	93.4	12.0	5235	19	AAV63022		D. immitis ankyrin		
9	93.4	12.0	5235	19	AAV63023		D. immitis ankyrin		
10	93.4	12.0	5235	21	AAA58195		D. immitis ankyrin		
11	93.4	12.0	5235	21	AAA58196		D. immitis ankyrin		

[illegible]

ALIGNMENTS

RESULT	1	
AA35852		
ID	AA35852	standard; DNA; 780 BP.
XX	XX	
XX	AA35852;	
XX		
DT	14-JUL-1999	(first entry)
XX		
XX	DNA encoding human gankyrin protein.	
DE		
XX		
KW	Gankyrin; apoptosis induction; diagnosis; treatment; cancer;	
KW	hepatocellular carcinoma; oncogenesis mechanism; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	W09918201-A1.	
XX		
PD	15-APR-1999.	
XX		
PF	02-OCT-1998;	98WO-JF04467.
XX		
PR	03-OCT-1997;	97JP-0286214.
XX		
PA	(FUJI/) FUJITA.	
XX		
PI	Fujita J;	
XX		
DR	WPI; 1999-277266/23.	
XX	P-PSDB; AAY02430.	
XX		
PT	Gankyrin polypeptides, useful for treatment and diagnosis of	
PT	cancers, e.g. hepatocellular carcinoma, and study of oncogen	
PT	mechanism	
XX		

PS Claim 5; page 68-69; 111pp; Japanese.

The specification describes human, murine and rat gankyrin DNA and polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic ability and apoptosis induction. The polypeptides and their antibodies can be used in the diagnosis and treatment of cancers, e.g. hepatocellular carcinoma, and study of oncogenesis mechanism. The present sequence encodes human gankyrin.

Sequence 780 BP: 222 A; 161 C; 221 G; 176 T; 0 other;

[illegible]

RESULT 2
AAT66425
ID AAT66425 standard; cDNA to mRNA; 1468 BP.
XX

AC	AAAT66425;	
XX	17-JUN-1997 (first entry)	
DT		
XX		
DE	Human P28 coding sequence.	
XX		
XX	Human; proteasome; P28; diagnosis; malignant tumour; ds.	
XX		
OS	Homo sapiens.	
XX		
XX	Location/Qualifiers	
FH	23..703	
FT	/*tag= a	
FT	/product= P28	
FT		
XX	JP09075085-A.	
PN		
XX	25-MAR-1997.	
PD		
XX		
XX	13-SEP-1995; 95JP-0235052.	
XX		
PF		
XX	13-SEP-1995; 95JP-0235052.	
XX		
PR	(SAGA) SAGAMI CHEM RES CENTRE.	
XX		
PA		
XX	WPI; 1997-239267/22.	
DR	P-PSDB; AAW15483.	
DR		
XX		
XX	Human 26S proteasome constituting component protein - useful in the	
PT	diagnosis of e.g. malignant tumour	
PT		
XX	Claim 4; Page 7-8; 9pp; Japanese.	
PS		
XX	This sequence encodes the human proteasome component protein p28.	
CC	The protein, p28, is useful for the diagnosis and treatment of	
CC	various diseases caused by proteasomes such as malignant tumour.	
CC		
XX	Sequence 1468 BP; 424 A; 268 C; 324 G; 452 T; 0 other;	
XX		
SO		

Query Match		90.5%;	Score 706;	DB 18;	Length 1468;
Best Local Similarity		100.0%;	Pred. No. 1.3e-218;		
Matches	706;	Conservative	0;	Mismatches	0; Indels 0; Gaps
Qy	75	aagtagtctgagcagcagcaaatggaggggtgtgtctaaacctaatggtctgcaacct	134		
Db	1	aagtagtctgagcagcaaatggaggggtgtgtctaaacctaatggtctgcaacct	60		
Qy	135	ggctacagcgggaagctgaaagattgaaggagagtattctctggcagataaatccctggc	194		
Db	61	ggctacagcgggaagctgaaagattgaaggagagtattctctggcagataaatccctggc	120		
Qy	195	tactagaactgaccagacagcagaaactgcattgcactgggcattgctcagctggacatac	254		
Db	121	tactagaactgaccagacagcagaaactgcattgcactgggcattgctcagctggacatac	180		
Qy	255	agaaattgtgaaatttttggcgaacttgagtgccagtgaaatgataaagacgatcagg	314		
Db	181	agaaattgtgaaatttttggcgaacttgagtgccagtgaaatgataaagacgatcagg	240		
Qy	315	ttgctctctcttcatttcggcgtctctgctgcccggagtgaagatgtgaaagccctct	374		
Db	241	ttggctctctcttcatttcggcgtctctgctgcccggagtgaagatgtgaaagccctct	300		
Qy	375	gggaaaagtgtcctaagtgaatgctgtcgaatcgaatcgaatcgaatcgaatcgaatcga	434		
Db	301	gggaaaagtgtcctaagtgaatgctgtcgaatcgaatcgaatcgaatcgaatcgaatcga	360		
Qy	435	agcttcgaaaacaggcatagatcgctgctcatgttactggaagcggggcgttaatccaga	494		
Db	361	agcttcgaaaacaggcatagatcgctgctcatgttactggaagcggggcgttaatccaga	420		
Qy	495	tgctaaggaccattatgatgggtacagcaatgcaccggcgagcgaagggtaacttgaa	554		

Db 421 |||||tgtaaggaccattatgagctacagcaatgcaccggcagcagcaagggttaacttgaa 480
QY 555 gatgattcatactctctgtactacaagaacatcccaacaacatcccaagacactgagggtaa 614
Db 481 gatgattcatactctctgtactacaagaacatcccaacaacatcccaagacactgagggtaa 540
QY 615 cactctctacacttagcctgtgatgagagagagtgaaagagcaaaactgctggttc 674
Db 541 cactctctacacttagcctgtgatgagagagagtgaaagagcaaaactgctggttc 600
QY 675 ccaaggagcaagtattacattgagaataaagaagaagaagacacccctgcaagtggccaa 734
Db 601 ccaaggagcaagtattacattgagaataaagaagaagaagacacccctgcaagtggccaa 660
QY 735 agtgggcctgggttaatactacaagagaatggtggaaggttaaaaaa 780
Db 661 agtgggcctgggttaatactacaagagaatggtggaaggttaaaaaa 706

RESULT 3

ID AAT66424 standard; cDNA to mRNA; 678 BP.
XX

AC AAT66424;
XX

DT 17-JUN-1997 (first entry)
XX

DE Human P28 coding sequence.
XX

KW Human; proteasome; P28; diagnosis; malignant tumour; ds.
XX

OS Homo sapiens.
XX

PN JP09075085-A.
XX

PD 25-MAR-1997.
XX

PF 13-SEP-1995; 95JP-0235052.
XX

PR 13-SEP-1995; 95JP-0235052.
XX

PA (SAGA) SAGAMI CHEM RES CENTRE.
XX

DR WPI; 1997-239267/22.
XX

DR P-PSDB; AAW15483.
XX

PT Human 26S proteasome constituting component protein - useful in the diagnosis of e.g. malignant tumour
XX

PS Claim 1; Page 6-7; 9pp; Japanese.
XX

CC This sequence encodes the human proteasome component protein P28.
XX

CC The protein, P28, is useful for the diagnosis and treatment of
XX

CC various diseases caused by proteasomes such as malignant tumour.
XX

SQ Sequence 678 BP; 201 A; 138 C; 183 G; 156 T; 0 other;
XX

Query Match 86.9%; Score 678; DB 18; Length 678;
Best Local Similarity 100.0%; Pred. No. 9.7e-210;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 atggaggggtgtgtctaaacctaatgtctgcaacctggcctacagcggaagctgaa 156
Db 1 atggaggggtgtgtctaaacctaatgtctgcaacctggcctacagcggaagctgaa 60
QY 157 gagttgaaggagagatctctggccgataaaatccctggctactagaaactgaccaggacagc 216
Db 61 gagttgaaggagagatctctggccgataaaatccctggctactagaaactgaccaggacagc 120
QY 217 aqaactgcattgcactggcctgctcagctggacatacagaaattgtgaattttgtg 276
|||||

Db 121 agaactgcattgcactggcctgctcagctggacatacagaaattgtgaattttgtg 180
QY 277 caacttgagtgccagtgaaatgataaagaacgatgaggttggtctcctcttcattatggc 336
Db 181 caacttgagtgccagtgaaatgataaagaacgatgaggttggtctcctcttcattatggc 240
QY 337 gcttctgctggccggatgagattgtaaaagccctctcgggaaaaaggtgctcaagtgaat 396
Db 241 gcttctgctggccggatgagattgtaaaagccctctcgggaaaaaggtgctcaagtgaat 300
QY 397 gctgtcaatcaaaaaggctgtactcctctacattatgcagcttcgaaaaacacaggcatgag 456
Db 301 gctgtcaatcaaaaaggctgtactcctctacattatgcagcttcgaaaaacacaggcatgag 360
QY 457 atcgtgtcatgttactggaagcgggcctaataccagatgctaaaggaccattatgaggt 516
Db 361 atcgtgtcatgttactggaagcgggcctaataccagatgctaaaggaccattatgaggt 420
QY 517 acagcaatgcaccgggcagcagcgaagggttaactggaagatgattcatatcctctgttac 576
Db 421 acagcaatgcaccgggcagcagcgaagggttaactggaagatgattcatatcctctgttac 480
QY 577 tacaagcatccacaacatccacagactgaggggttaacactcctctacacttagcctgt 636
Db 481 tacaagcatccacaacatccacagactgaggggttaacactcctctacacttagcctgt 540
QY 637 gatgaggagagtggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 696
Db 541 gatgaggagagtggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 600
QY 697 gagaataaagaagaagaacacccctgcaagtggccaaaggtggcctgggtttaatactc 756
Db 601 gagaataaagaagaagaacacccctgcaagtggccaaaggtggcctgggtttaatactc 660
QY 757 aagagaatggtggaaggt 774
Db 661 aagagaatggtggaaggt 678

RESULT 4

AA35854
ID AAX35854 standard; DNA; 696 BP.
XX

AC AAX35854;
XX

DT 14-JUL-1999 (first entry)
XX

DE DNA encoding rat gankyrin protein.
XX

KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism; ss.
XX

OS Rattus sp.
XX

PN W09918201-A1.
XX

PD 15-APR-1999.
XX

PF 02-OCT-1998; 98WO-JP04467.
XX

PR 03-OCT-1997; 97JP-0286214.
XX

PA (FUJI/) FUJITA.
XX

PI Fujita J;
XX

DR WPI; 1999-277266/23.
XX

DR P-PSDB; AAY02432.
XX

PT Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX

```
PS Claim 15; Page 75-76; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence encodes rat gankyrin.
XX
SQ Sequence 696 BP; 211 A; 140 C; 183 G; 162 T; 0 other;

Query Match 76.0%; Score 593; DB 20; Length 696;
Best Local Similarity 91.0%; Pred. No. 3.7e-182;
Matches 626; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 97 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggaa 156
Db 1 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacacgggaagctggat 60

Qy 157 gagttgaagagagatattctggccgataaatacctggctactagaactgaccaggacagc 216
Db 61 gagttgaaggaagcatttggctgataagctctggtccactagaactgacaggacagc 120

Qy 217 agaactgattgcactggcagctgcagctggacacacagaaattgtgaattttgtg 276
Db 121 agaacagcattgcactggcagctgcagctggacacacagaaattgtgaattttgtg 180

Qy 277 caacttgagtgccagtgaaatgataaagccttgcctgactagaactgttgaaattttgtg 336
Db 181 caacttgagtgccagtgaaatgataaagccttgcctgactagaactgttgaaattttgtg 240

Qy 337 gcttctgctggccgggagtgagattgtaaaagcccttctgggaaagtgctcgaagtgaat 396
Db 241 gcttctgctggccgggagtgagattgtaaaagcccttctgggaaagtgctcgaagtgaat 300

Qy 397 gctgtcaatacaaaatggctgtactcccttaccattatcagcttcgaaaaaacaggcatgag 456
Db 301 tctgtcaatacaaaagcgtgcactcccttaccattatcagcttcgaaaaaataggcatgag 360

Qy 457 atcgtgtcattgttactgggaagggcggttaatacagatcgaagcattatgaggt 516
Db 361 attctgtattgttactagaaggtggggttaacccagatgocgaagaccattacagtgct 420

Qy 517 acagcaatgcacgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
Db 421 acagcaatgcacgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480

Qy 577 tacaaagcatccacaaacatccaaagacactgagggtaaacactcctctacacttagcctgt 636
Db 481 tacaaagcatccacaaacatccaaagacactgagggtaaacactcctctacacttagcctgt 540

Qy 637 gatgaggagagtggaagaagcaaaactgctgtgtgtcccaaggagcaagtattacatt 696
Db 541 gatgaggagagtggaagaagcaaaactgctgtgtgtgtgacccaaggagcaagtattacatt 600

Qy 697 gagaataagaagaagaacacacccctgcaagtgcccaagtggtgacctgggttatactc 756
Db 601 gaaataaggaagaagaacacacccctgcaagtcgccaaggggacctgggttatactc 660

Qy 757 aagagaatggtggaagggttaa 777
Db 661 aagaagaatgcagaaagttaa 681

RESULT 5
AA335853
ID AAX35853 standard; DNA; 696 BP.
XX
AC AAX35853;
XX
DT 14-JUL-1999 (first entry)
XX
```

```
DE DNA encoding mouse gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism; ss.
XX
OS Mus sp.
XX
PN WO9518201-A1.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR P-PSDB; AAY02431.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 10; Page 71-73; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence encodes mouse gankyrin.
XX
SQ Sequence 696 BP; 205 A; 142 C; 180 G; 168 T; 1 other;

Query Match 74.9%; Score 584; DB 20; Length 696;
Best Local Similarity 91.0%; Pred. No. 3e-179;
Matches 620; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 97 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggaa 156
Db 1 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggat 60

Qy 157 gagttgaagagagatattctggccgataaatacctggctactagaactgaccaggacagc 216
Db 61 gagttgaagagagcagcatttggctgataaatacctggctactagaactgacaggacagc 120

Qy 217 agaactgattgcactggcagctgcagctggacacacagaaattgtgaattttgtg 276
Db 121 agaacagcttgcactggcagctgcagctggacacacagaaattgtgaattttgtg 180

Qy 277 caacttgagtgccagtgaaatgataaagccttgcctgactagaactgttgaaattttgtg 336
Db 181 caacttgagtgccagtgaaatgataaagccttgcctgactagaactgttgaaattttgtg 240

Qy 337 gcttctgctggccgggagtgagattgtaaaagcccttctgggaaagtgctcgaagtgaat 396
Db 241 gcttctgctggccgggagtgagattgtaaaagcccttctgggaaagtgctcgaagtgaat 300

Qy 397 gctgtcaatacaaaatggctgtactcccttaccattatcagcttcgaaaaaacaggcatgag 456
Db 301 tctgtcaatacaaaagcgtgcactcccttaccattatcagcttcgaaaaaataggcatgag 360

Qy 457 atcgtgtcattgttactgggaagggcggttaatacagatcgaagcattatgaggt 516
Db 361 attctgtattgttactagaaggtggggttaacccagatgocgaagaccattacagtgct 420

Qy 517 acagcaatgcacgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
Db 421 acagcaatgcacgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
```


Qy	662	aactcgtggttcccaaggaacgaagtatttacataggataaagaagaacaccoc	721
Db	3353	GCACCTCTCCTTCATTATAAGCAAAATGCGAATGTGTAAGCAAAGCTTGCTTTACACCAC	3294
Qy	722	tccaagtggcc	732
Db	3293	TTCATCTTGGCC	3283
RESULT	10		
AAA58195			
ID	AAA58195 standard; cDNA; 5235 BP.		
XX	AAA58195;		
XX	AC		
XX	AC		
DT	23-OCT-2000 (first entry)		
XX			
DE	D. immitis ankryrin cDNA coding sequence, nDiAnk5235.		
XX			
KW	Ankryrin; parasitic helminth; filariid nematode; heartworm disease;		
KW	elephantiasis; hydrocele; vaccine; antibody; antihelminthic; ss.		
CS	Dirofilaria immitis.		
XX			
PN	US6063599-A.		
PD	16-MAY-2000.		
PF	24-APR-1998; 98US-0065474..		
PR	24-APR-1997; 97US-0847429..		
PA	{HESK-} HESKA CORP.		
PI	Blehm ES, Tang L;		
DR	WPI; 2000-375493/32.		
P-	PSDB; AAB11589.		
XX			
PT	New Dirofilaria and Brugia ankryrin proteins and nucleic acid encoding		
PT	them, useful for treating and protecting animals from diseases caused		
PT	by parasitic helminths, e.g. heartworm disease, elephantiasis or		
PT	hydrocele -		
XX			
PS	Example 1; Column 133-138; 120pp; English.		
CC	The invention relates to ankryrin proteins and nucleic acids from the		
CC	parasitic helminths Dirofilaria immitis and Brugia malayi. It also		
CC	relates to antibodies raised against such ankryrin proteins and to		
CC	compounds that inhibit Dirofilaria or Brugia ankryrin function.		
CC	Dirofilaria ankryrin cDNAs were isolated from a D. immitis 48 hour		
CC	L3 cDNA library using PCR primers based on the sequence of the El		
CC	ankryrin from Onchocerca volvulus and the Caenorhabditis elegans ankryrin		
CC	UNC-44 genes. Brugia ankryrin cDNAs were isolated from a B. malayi adult		
CC	female cDNA library using D. immitis ankryrin and C. elegans UNC-44 PCR		
CC	primers. Dirofilaria or Brugia ankryrin proteins and nucleic acids		
CC	represent novel targets for anti-helminthic vaccines and drugs. Ankryrin		
CC	nucleic acid molecules, proteins, vaccines and compositions are useful		
CC	for protecting animals, particularly dogs, from diseases caused by		
CC	parasitic helminths (e.g., heartworm disease, elephantiasis or		
CC	hydrocele), as well as for treating the infection. The ankryrin nucleic		
CC	acid molecules, proteins, vaccines and compositions of the invention are		
CC	especially useful in treating and preventing infections caused filariid		
CC	nematodes (e.g., D. immitis and B. malayi), and ascariid, capillarid,		
CC	strongylid, strongyloides, trichostrongyle, or trichurid nematodes, and		
CC	are also useful against cestodes and trematodes. The therapeutic		
CC	compositions may be administered to mammals, including dogs, cats,		
CC	humans, ferrets, horses, cattle, sheep, and other pets; economic food		
CC	animals; or zoo animals. The ankryrin nucleic acid molecules, proteins and		
CC	compounds may also be used as diagnostic reagents to detect infection by		
CC	parasitic helminths. Prior art anti-helminthic drugs require repeated		
CC	administration, which often leads to the development of resistant		


```
XX The present sequence encodes a full length Dirofilaria immitis ankyrin
CC protein. The ankyrin protein, or anti-ankyrin antibodies, may be used to
CC protect an animal from disease caused by a parasitic helminth,
CC especially where the disease is heartworm disease, elephantiasis or
CC hydrocele.
XX
SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;

Query Match      12.0%; Score 93.4; DB 19; Length 5503;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgacacagacagagaactgcattgcaactggcgatgct 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 atggagcacaggtggatgctgctgctggaactacaactccactgcacattgcatcac 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 cagctggacacagaaattgtgaattttgtgcaacttgagtgccagtgaaatgata 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1513 gtcttgtaataccgacatcgctatttctgctgaggttaatacaccacaaatgctg 1572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 aagacgatgcaggttgctctctctctctctctctctctctctctctctctctctct 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1573 ccacaagagatcttatactctctctctctctctctctctctctctctctctctctct 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 taaaagccctctgggaaaaaggtgctcaagtgaatgctgcaatcaaaaatggctgactc 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1633 cagcaatattgattgattggaaccgacagacactgctcaccgaaaggggttttaccgc 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 ccttacattatgcagctctgaaaaaacaggcatgagatcgctgctgctgctgctgctgctg 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1693 cgttgatttagctgtaagtattgcaatttgcggtcggaataatcatctgctgagaacgag 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 gggctaataccagatgctgaaggaccattatgaggtcctacagcaatgcaccgggagcagcca 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1753 gaacacgggttgacattggaaggcaagaatacagtaaacctctgcatgtgagcgacatt 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 agggtaacttgaaagattcattctctctgctgactacaaagacatccacaaatccaag 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1813 acaataacgacaaggtagcatttcttctctgagaataatggtgcttctgacatgcgcgtg 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 602 acactgaggttaacactcctctacacttagcctgctgagggagagagagtggaagaagcaa 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1873 ccaagaatgggtacactcctctacattatgctccgaggaagaatacagatggtattgcta 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 aactgctggtgtcccaaggcgaagtatttacctgagaaataaagaagaacacacccc 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1933 gcactctcctcattataaggcaaatgcgaatgctgaaagcaaaagctggctttacaccac 1992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 722 tgcaagtggcc 732
   ||| ||| |||
Db 1993 ttcatttgcc 2003
   ||| ||| |||

RESULT 13
AAV63020
ID AAV63020 standard; cDNA; 5503 BP.
XX
AC AAV63020;
XX
XX 15-JAN-1999 (first entry)
DT
DE
DE D. immitis ankyrin nDiAnk5503 cDNA.
XX
KW Ankyrin; helminth; parasite; vaccine; infection;
KW passive immunogen; cytotoxic agent; ss.
XX
XX Dirofilaria immitis.
OS
XX Key Location/Qualifiers
PH 51..5288
FT CDS
FT /*tag= a
```

```
FT XX /product= "nDiAnk5503"
PN XX US5827692-A.
XX XX 27-OCT-1998.
PD XX
PF XX 24-APR-1997; 97US-0847429.
XX XX
PR XX 24-APR-1997; 97US-0847429.
XX XX
PA (HESK-) HESKA CORP.
XX XX
PI Blehm ES, Tang L;
XX XX
XX WPI: 1998-593992/50.
XX P-PSDB; AAW76776.
XX
PT Nucleic acids encoding ankyrins from helminth parasites - useful for
PT recombinant production of the proteins for use as vaccines and
PT treatments against helminth infection
PT
XX Claim 1; Column 95-108; 84pp; English.
PS
XX AAV62996-V63027 encode ankyrin proteins isolated from the helminth
CC parasites Dirofilaria immitis and Brugia malayi. The nucleic acids and
CC recombinant products are useful for the recombinant production of the
CC ankyrin polypeptides. These proteins can then be used as vaccines
CC against parasitic helminth, e.g. D. immitis or B. malayi. They can also
CC be used for therapy after infection, and to raise antibodies against
CC use in therapeutics, as passive immunogens, or as therapeutics against
CC helminths on conjugation to cytotoxic agents. The nucleic acids contained
CC in viruses, may also be used as viral vaccines, and the nucleic acids
CC themselves or in vectors may be used as genetic vaccines.
XX
SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;

Query Match      12.0%; Score 93.4; DB 19; Length 5503;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgacacagacagagaactgcattgcaactggcgatgct 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 atggagcacaggtggatgctgctgctggaactacaactccactgcacattgcatcac 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 cagctggacacagaaattgtgaattttgtgcaacttgagtgccagtgaaatgata 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1513 gtcttgtaataccgacatcgctatttctgctgaggttaatacaccacaaatgctg 1572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 aagacgatgcaggttgctctctctctctctctctctctctctctctctctctctct 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1573 ccacaagagatcttatactctctctctctctctctctctctctctctctctctctct 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 taaaagccctctgggaaaaaggtgctcaagtgaatgctgcaatcaaaaatggctgactc 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1633 cagcaatattgattgattggaaccgacagacactgctcaccgaaaggggttttaccgc 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 ccttacattatgcagctctgaaaaaacaggcatgagatcgctgctgctgctgctgctgctg 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1693 cgttgatttagctgtaagtattgcaatttgcggtcggaataatcatctgctgagaacgag 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 gggctaataccagatgctgaaggaccattatgaggtcctacagcaatgcaccgggagcagcca 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1753 gaacacgggttgacattggaaggcaagaatacagtaaacctctgcatgtgagcgacatt 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 agggtaacttgaaagattcattctctgctgactacaaagacatccacaaatccaag 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1813 acaataacgacaaggtagcatttcttctctgagaataatggtgcttctgacatgcgcgtg 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 602 acactgaggttaacactcctctacacttagcctgctgagggagagagagtggaagaagcaa 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1873 ccaagaatgggtacactcctctacattatgctccgaggaagaatacagatggtattgcta 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:46:08 ; Search time 80.36 Seconds
(without alignments)
1837.515 Million cell updates/sec

Title: US-09-509-775-1

Perfect score: 780

Sequence: 1 tggtagacttaacgctg.....gaatggtggaagttaaca 780

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PT05_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.4	12.0	1056	3	US-09-065-474-138
2	93.4	12.0	1056	3	US-09-065-474-140
3	93.4	12.0	5235	1	US-09-031-485-35
4	93.4	12.0	5235	1	US-09-031-485-36
5	93.4	12.0	5235	1	US-08-847-429A-35
6	93.4	12.0	5235	1	US-08-847-429A-36
7	93.4	12.0	5235	3	US-09-065-474-35
8	93.4	12.0	5235	3	US-09-065-474-36
9	93.4	12.0	5503	1	US-09-031-485-32
10	93.4	12.0	5503	1	US-09-031-485-34
11	93.4	12.0	5503	1	US-08-847-429A-32
12	93.4	12.0	5503	1	US-08-847-429A-34
13	93.4	12.0	5503	3	US-09-065-474-32
14	93.4	12.0	5503	3	US-09-065-474-34
15	88.4	11.3	3454	4	US-09-082-059-1
16	86.6	11.1	911	1	US-09-031-485-22
17	86.6	11.1	911	1	US-09-031-485-24
18	86.6	11.1	911	1	US-08-847-429A-22
19	86.6	11.1	911	1	US-08-847-429A-24
20	86.6	11.1	911	3	US-09-065-474-22
21	86.6	11.1	911	3	US-09-065-474-24
22	85.6	11.0	909	1	US-09-031-485-25
23	85.6	11.0	909	1	US-09-031-485-26
24	85.6	11.0	909	1	US-08-847-429A-25
25	85.6	11.0	909	1	US-08-847-429A-26
26	85.6	11.0	909	3	US-09-065-474-25
27	85.6	11.0	909	3	US-09-065-474-26

28	78.6	10.1	908	1	US-09-031-485-37	Sequence 37, Appl
c 29	78.6	10.1	908	1	US-09-031-485-39	Sequence 39, Appl
c 30	78.6	10.1	908	1	US-08-847-429A-37	Sequence 37, Appl
c 31	78.6	10.1	908	1	US-08-847-429A-39	Sequence 39, Appl
c 32	78.6	10.1	908	3	US-09-065-474-37	Sequence 37, Appl
c 33	78.6	10.1	908	3	US-09-065-474-39	Sequence 39, Appl
c 34	77.6	9.9	906	1	US-09-031-485-40	Sequence 40, Appl
c 35	77.6	9.9	906	1	US-09-031-485-41	Sequence 41, Appl
c 36	77.6	9.9	906	1	US-08-847-429A-40	Sequence 40, Appl
c 37	77.6	9.9	906	1	US-08-847-429A-41	Sequence 41, Appl
c 38	77.6	9.9	906	3	US-09-065-474-40	Sequence 40, Appl
c 39	77.6	9.9	906	3	US-09-065-474-41	Sequence 41, Appl
c 40	74.4	9.5	573	1	US-09-031-485-19	Sequence 19, Appl
c 41	74.4	9.5	573	1	US-09-031-485-21	Sequence 21, Appl
c 42	74.4	9.5	573	1	US-08-847-429A-19	Sequence 19, Appl
c 43	74.4	9.5	573	1	US-08-847-429A-21	Sequence 21, Appl
c 44	74.4	9.5	573	3	US-09-065-474-19	Sequence 19, Appl
c 45	74.4	9.5	573	3	US-09-065-474-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-065-474-138
; Sequence 138, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1056
US-09-065-474-138

Query Match 12.0%; Score 93.4; DB 3; Length 1056;
Best Local Similarity 48.1%; Pred. No. 8.3e-21;

Query Match	12.08;	Score 93.4;	DB 1;	Length 5503;
Best Local Similarity	48.1#;	Pred. No. 2.2e-20;		
Matches 265;	Conservative 0;	Mismatches 286;	Indels 0;	Gaps 0;
QY	182	ataaatccctggctactagaactgaccagacagacagaactgcattcacctgggcattgct	241	
Db	1453	ATGGAGACACAGGTGGATGCTGCTCTGCTGAACATACAACTCCACTGTCACATTCGATCAC	1512	
QY	242	cagctggacatacagaattgttgaattttttgttgcaactctggagtgccagtggaatgata	301	
Db	1513	GTCTTGTGTAATACCGACATCGTCAATTTGTGTCTGCAGGCTAAATGCATCACCATAATGCTG	1572	
QY	302	agacgatgcaggttgctccctctccatctcattcggctctctgctggcgggatagattg	361	
Db	1573	CCACAAGAGATCTTTATACCTCTTCAATGTGTCGCAAGAGGGGCAAGGAAGTGG	1632	
QY	362	taaaagccctctgggaaaaagggtgctcaagtgaatgctgtcaatcaaaatgctgtgactc	421	
Db	1633	CAGCAATATTGATGGATCATGTGAACCGCAACAGACACTGCTCAGCAAAAAGSGTTTACGC	1692	
QY	422	ccctacattatgcagcttcgaaaaaacaggcatgagatgcgtgctgcttactggaagcg	481	
Db	1693	CGTTGCATTTAGCTGCTTAAGTATGGCAATTTGCCGGTTCGCGAAATCATTTGTGAACGAG	1752	
QY	482	gggctaatccagatgctaaaggaccattatgaggctacagcaatgcacggcgacagcca	541	
Db	1753	GAACACCGGTTGACATTTGAAGGCAAGAATCAGGTAACACCTCTCATGTAGCGCACATT	1812	
QY	542	agggtaactggaatgattcattatcctctgttactacaagcatccacaaatccaag	601	
Db	1813	ACAATACGACAGGTAGCATTTGTTACTCTTAGAAAATGGTGTCTTCGCATATGCCGCTG	1872	
QY	602	acactgagggtaacactccctctacacttagcctgtgatgaggagagagtggaagaagcaa	661	

Query Match	12.0%	Score 93.4	DB 1	Length 5503
Best Local Similarity	48.1%	Pred. No. 2.2e-20		
Matches 265	Conservative	0	Mismatches 286	Indels 0
Gaps	0			
Qy 182	ataaatccccgtgctactagaaactaccagagacagcagagaactgcattgcactgggcgatcgt	241		
Db 4051	ATGGAGCAGACGTTGGATGCTGCTGCTGTGAACATACAAACTCCACTGGCACATTCATCAC	3992		
Qy 242	cagctgcacatacagaanaattgttgaaattttgttgcaacttggagtgccagtgaaatgata	301		
Db 3991	GTCTTGTAATACCGACATCGTCATTTTGTGTGTCAGGCTAATGCATCACCAAAATGCTG	3932		
Qy 302	aagacgatgcaggttgctctctctctctcatattgcggcttctgctgcccggatgagattg	361		
Db 3931	CCACAAGAGATCTTTATACCTCTCTCATATGTGTCGAAGAGGGGGCAAGAGAAGTGG	3872		

OM of: US-09-509-775-1 to: PIR_68:* out_format : pfs

Date: Aug 13, 2001 8:34 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.n2p.model -DEV=xlp
-O=/cnp2_1/USPTO.spool/US09509775/runat_13082001_074528_13056/app_query.fasta_1.848
-DB=PIR_68 -QPM=fastan -SUFFIX=tr -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09509775 -CGN1_1_64 -NCPU=6 -TCPU=3
-LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-509-775-1

Query length: 780

Database: PIR_68*

Database sequences: 219241

Database length: 76174552

Search time (sec): 92.960000

score_list:	Strd	Orig	ZScore	Escore	Len	Documentation
Sequence	+	338.00	560.20	1.3e-23	247	! probable ankyrin [imported] - A
pir2:D84448	+	312.50	500.61	3.8e-21	1765	! ankyrin 3, splice form 2 - mou
pir2:T42714	+	312.50	499.82	3.9e-21	1940	! ankyrin 3, splice form 3 - mou
pir2:T42715	+	312.50	499.80	3.9e-21	1943	! ankyrin 3, splice form 1 - mou
pir2:T42716	+	312.50	499.73	3.9e-21	1961	! ankyrin 3, splice form 4 - mou
pir2:T50984	+	310.00	513.29	5.6e-21	237	! related to 26s proteasome subu
pir2:A55575	+	310.00	488.75	7.1e-21	4377	! ankyrin 3, long splice form -
pir2:T33631	+	305.00	496.54	1.8e-20	636	! hypothetical protein F4069.1 -
pir2:T37431	+	294.00	462.66	2.2e-19	3924	! ankyrin 2, neuronal long splic
pir2:T57697	+	291.00	481.54	3.4e-19	228	! hypothetical protein YG232w -
pir2:B35049	+	290.00	462.21	5.0e-19	1856	! ankyrin 1, erythrocyte splice
pir2:A35049	+	290.00	462.10	5.0e-19	1880	! ankyrin 1, erythrocyte splice
pir1:SIJHUK	+	290.00	462.10	5.0e-19	1881	! ankyrin 1, erythrocyte splice
pir2:A57282	+	288.50	460.00	6.9e-19	1786	! ankyrin-related protein unc-44
pir2:T15346	+	288.50	459.87	7.0e-19	1815	! elegans ankyrin-related unc-44
pir2:T15344	+	288.50	459.63	7.0e-19	1867	! ankyrin-related unc-44 - Caen
pir2:T15347	+	288.50	458.89	7.0e-19	2039	! ankyrin-related unc-44 - Caen
pir2:T3771	+	287.50	458.03	8.7e-19	1848	! ankyrin, erythrocyte - mouse
pir2:T49502	+	287.50	457.97	8.7e-19	1862	! ankyrin - mouse
pir2:T13940	+	276.50	440.95	9.2e-18	1549	! ankyrin - fruit fly (Drosophi
pir2:T39032	+	273.50	451.79	1.5e-17	234	! hypothetical ankyrin-like prote
pir2:TJG0197	+	270.50	436.22	3.2e-17	815	! myosin-light-chain-phosphatase
pir2:S68418	+	269.50	436.34	3.9e-17	658	! protein phosphatase 1M chain M
pir2:A55142	+	268.50	431.09	5.0e-17	1004	! myosin-light-chain-phosphatase
pir2:T30255	+	263.00	421.34	1.7e-16	1062	! inversin - mouse
pir2:T14151	+	256.00	409.52	4.7e-16	1062	! inv protein - mouse
pir2:T46445	+	248.50	405.14	3.6e-15	397	! hypothetical protein DKFZp34B2
pir2:T151527	+	247.50	392.85	4.9e-15	1403	! alpha-latrotoxin precursor - h
pir1:T37275	+	247.50	391.87	5.4e-15	1421	! death-associated protein kinase
pir2:S30355	+	243.50	386.03	1.2e-14	1411	! alpha-latroinsectotoxin precu
pir2:T18184	+	239.50	390.59	2.5e-14	368	! ankyrin repeat protein A682L -
pir2:A57291	+	239.00	390.95	2.7e-14	319	! cytokine inducible nuclear prot
pir2:T46507	+	239.00	386.26	2.9e-14	557	! hypothetical protein DKFZp586M2
pir2:T42691	+	237.50	380.78	4.1e-14	791	! hypothetical protein DKFZp34D2
pir2:T12503	+	237.50	380.16	4.1e-14	851	! hypothetical protein DKFZp34A1
pir2:TJ01729	+	236.50	384.04	4.8e-14	439	! ankyrin-repeat protein - Arabid
pir2:T43458	+	234.50	373.48	7.9e-14	1031	! hypothetical protein DKFZp434E
pir1:H71274	+	233.00	371.78	1.1e-13	934	! probable ankyrin - syphilis spi
pir2:T21884	+	232.00	366.70	1.4e-13	1398	! hypothetical protein F36H1.2 -
pir2:T12952	+	229.00	363.01	2.6e-13	1188	! hypothetical protein C29E6.2 -
pir2:T19006	+	228.00	362.64	3.2e-13	1016	! ankyrin related protein C06C3

pir2:T30201 + 228.00 355.57 3.5e-13 2352 ! Notch homolog protein - sea
pir2:S30385 + 226.50 360.23 4.5e-13 1001 ! G9a protein - human
pir2:B47169 + 225.00 367.21 5.7e-13 323 ! ankyrin-like repeat protein
pir2:A42024 + 224.50 357.75 6.8e-13 900 ! transcription factor NF-kapp

seq_name: pir2:D84448

seq_documentation_block:

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84448
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84448
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: GB:AE002093; NID:g4335756; PIDN:AAD17433.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g03430
A:Map position: 2

alignment_scores:

Quality: 338.00 Length: 219
Ratio: 2.315 Gaps: 6
Percent Similarity: 66.667 Percent Identity: 41.096

alignment_block:

US-09-509-775-1 x D84448

Align seg 1/1 to: D84448 from: 1 to: 247

160 TTCAAGGAGAGATTCTGCGCCCATTAATCCCTGGCTACTAGAACTGACCA 209
|||||:|||||: |||: ||||||| ||| :|||:
31 LeuSerGluGluGlnLeuSer...LysSerLeuAsnPheArg...AsnG1 45
210 GGACAGCAGCAACTGCATCTGCACCTGGCGCTCAGCTGCACATACAGAAA 259
|||||:|||||: |||: ||||||| |||: |||:
45 uaspGlyArgSerLeuLeuHisValalaSerPheGlyHisSerGlnI 62
260 TTCTTCAATTTTGTGTG.....CAACTTGGAGTGCACGTGAATGAT 300
|||||:|||||: |||: ||||||| |||: |||:
62 leValLysLeuLeuSerSerSerAspGluAlaLysThrValIleAsnSer 78
301 AAAGACGATGAGTGGTCTCTCTTCATATTCGGCTTCTGCTGCGCG 350
|||||:|||||: |||: ||||||| |||: |||:
79 LysAspAspGluGlyTrpAlaProLeuHisSerAlaAlaSerIleGlyAs 95
351 GCATGAGATTGTAAAGCCCTTCTGGGAAAGGCTCAAGTGAATGCTG 400
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
95 nalaGluLeuValGluValLeuLeuThrGlyAlaAspValAsnAlaL 112
401 TCAATCAAAATGGCTGTACTCCCTTCATATTGACATTCGAAACAGG 450
|||||:|||||: |||: ||||||| |||: |||:
112 ysAsnAsnGlyGlyArgThrAlaLeuHisTyAlaAlaSerLysGlyArg 128
451 CATGAGATCGCTGTCTTATTCATGAGCGGGGCTTAATCCAGATGCTAA 500
|||||:|||||: |||: ||||||| |||: |||:
129 LeuGluIleAlaGlnLeuLeuLeuThrHisGlyAlaLysIleAsnIleth 145
501 GGACCATATTGAGGCTACAGCAATGCACCGGCGCAGCAGCAAGGTAAC 550
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
145 rAspLysValGlyCysThrProLeuHisArgAlaAlaSerValGlyLysL 162
551 TGAAGATGATTCATCTCTCTGTGTACTACAAAGACATCCCAACATCAA 600
|||||:|||||: |||: ||||||| |||: |||:
162 euGluValCysGluPheLeuIleGluGluGlyAlaGluIleAspAlaThr 178

220 ACTGCATTGGCACTGGGCATGCTCAGCTGGACATACAGAAATTGTTGAATT 269

```

517 ThrProLeuHisLeuAlaAlaArgGluGlyHisGluAspValAlaAlaPh 533
270 TTTGTTCCAACTTGGAGTCCAGTGATGATAAAGACGATGCAGGTGGT 319
533 eLeuLeuAspHisGlyAlaSerLeuSerIleThrThrLysLysGlyPheT 550
320 CTCTCTTCATATTGCGGCTTCTGCTGCGGGATGAGATTGTAAGGCC 369
550 hrProLeuHisValAlaAlaLysTyrglyLysLeuGluValAlaSerLeu 566
370 CTCTCTGGGAAAA 381
567 LeuLeuGlnLysSerAlaSerProAspAlaAlaGlyLysSerGlyLeuTh 583
381 ..... 381
583 rProLeuHisValAlaAlaHisTyraSpAsnGlnLysValAlaLeuLeuL 600
382 .....GGTCTCAAGTGAATGCTCTCAATCAAAATGGCTGACT 420
600 euLeuaspGlnGlyAlaSerProHisAlaAlaLysAsnGlyTyThr 616
421 CCCTTACATTATGACGCTTCGAAACACGATGATGATGATGATGATGAT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGAAGCGGGCTTAATCCAGATGCTAAGGACCATTTATGAGCTACAG 520
633 uLeuGluTyrglyAlaAlaLysAlaAlaValThrArgGlnGlyIleAla 650
521 CAATGCACCGGGGAGCAGCAAGGTAACCTGAAGATGATTTCATCTT 570
650 erValHisLeuAlaAlaGlnGlyHisValaspMetValSerLeuLeu 666
571 CTGACTACAAAGCATCCACAAACATCCAAAGACATGAGGGTAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCCTACACTTACCTGATGATGATGATGATGATGATGATGATGATGATG 670
683 oLeuHisLeuAlaAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCAAGGAGCAAGTATTACATTTCAGATAAAGAAAGAAAGACACC 720
700 alAsnGlnGlyAlaHisValaspAlaGlnThrLysMetGlyTyThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719
```

seq_name: pir2:T42713

```

seq_documentation_block:
ankyrin 3, splice form 1 - mouse
C.Species: Mus musculus (house mouse)
C.Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C.Accession: T42713
R.Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, J.
J. Cell Biol. 130, 313-330, 1995
A.Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A.Reference number: Z22237; MUID:95340633
A.Accession: T42713
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-1943 <PET>
A.Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1
A.Experimental source: strain C57BL/6J; kidney
C.Genetics:
A.Gene: Ank3
A.Map position: 10
A.Introns: 855/1
```

C:Function:
A:Description: supposed to play an important role in the polarized distribution of ma
A:Note: major kidney ankyrin
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

alignment_scores:
Quality: 312.50 Length: 203
Ratio: 2.500 Gaps: 1
Percent Similarity: 61.576 Percent Identity: 35.468
alignment_block:
US-09-509-775-1 x T42713 ..

Align seg 1/1 to: T42713 from: 1 to: 1943

```

220 ACTGCATTGCATGGGCGATCCTCAGCTGGACATACAGAAATTTGTAATT 269
517 ThrProLeuHisLeuAlaAlaArgGluGlyHisGluAspValAlaAlaPh 533
270 TTTGTTCCAACTTGGAGTCCAGTGATGATAAAGACGATGCAGGTGGT 319
533 eLeuLeuAspHisGlyAlaSerLeuSerIleThrThrLysLysGlyPheT 550
320 CTCTCTTCATATTGCGGCTTCTGCTGCGGGATGAGATTGTAAGGCC 369
550 hrProLeuHisValAlaAlaLysTyrglyLysLeuGluValAlaSerLeu 566
370 CTCTCTGGGAAAA 381
567 LeuLeuGlnLysSerAlaSerProAspAlaAlaGlyLysSerGlyLeuTh 583
381 ..... 381
583 rProLeuHisValAlaAlaHisTyraSpAsnGlnLysValAlaLeuLeuL 600
382 .....GGTCTCAAGTGAATGCTCTCAATCAAAATGGCTGACT 420
600 euLeuaspGlnGlyAlaSerProHisAlaAlaLysAsnGlyTyThr 616
421 CCCTTACATTATGACGCTTCGAAACACGATGATGATGATGATGATGAT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGAAGCGGGCTTAATCCAGATGCTAAGGACCATTTATGAGCTACAG 520
633 uLeuGluTyrglyAlaAlaLysAlaAlaValThrArgGlnGlyIleAla 650
521 CAATGCACCGGGGAGCAGCAAGGTAACCTGAAGATGATTTCATCTT 570
650 erValHisLeuAlaAlaGlnGlyHisValaspMetValSerLeuLeu 666
571 CTGACTACAAAGCATCCACAAACATCCAAAGACATGAGGGTAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCCTACACTTACCTGATGATGATGATGATGATGATGATGATGATGATG 670
683 oLeuHisLeuAlaAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCAAGGAGCAAGTATTACATTTCAGATAAAGAAAGAAAGACACC 720
700 alAsnGlnGlyAlaHisValaspAlaGlnThrLysMetGlyTyThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719
```

seq_name: pir2:T42716

seq_documentation_block:
ankyrin 3, splice form 4 - mouse

550 TTGAAGATGATTCATATCTCTTCTACTACAAAGCATCCACAAACATCCA 599
 ||| : : : : : ||| : : : : :
 582 LeuAspIleValLysLeuProArgGlySerProHisSerPr 598
 600 AGACATGAGGTTACACTCTCTACACTTACCTGTGATGAGGAGAG 649
 : : : : : ||| : : : : : ||| : : : : :
 598 oAlaTrpAsnGlyThrProLeuHisIleAlaAlaLysGlnAsnGlnV 615
 650 TGAAGAAGCAAACTGCTGTGTCTCCCAAGGAGCAAGTATTATTCATGAG 699
 ||| : : : : : ||| : : : : : ||| : : : : :
 615 aGluValAlaArgSerLeuLeuGlnThrGlySerAlaAsnAlaGlu 631
 700 AATAAAGAGAAAGACACCCCTGCAAGTGCC.....AAAGTGCC... 741
 : : : : : ||| : : : : : ||| : : : : :
 632 SerValGlnGlyValThrProLeuHisLeuAlaGlnGluGlyHisAl 648
 742CTGGGTTTAACTCAAGAGCAAGTGGTGAAGGT 774
 : : : : : ||| : : : : : ||| : : : : :
 648 aGluMetValAlaLeuLeuSerLysGlnAlaAsnGly 661

seq_name: pir1:A35049

seq_documentation_block:

N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
 N:Contains: ankyrin 2.2, erythrocyte
 C:Species: Homo sapiens (man)
 C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
 C:Accession: A35049
 R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
 Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
 A:Title: cDNA sequence for human erythrocyte ankyrin.
 A:Reference number: A35049; MUID:90175370

A:Accession: A35049

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1880 <LAN>

A:Cross-references: GB:M28880

C:Genetics:

A:Gene: GDB:ANK1; ANK

A:Cross-references: GDB:118737; OMIM:182900

A:Map position: 8p11.2-8p11.2

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing; cytoskeleton

F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>

F:2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>

F:44-76/Domain: ankyrin repeat homology <AN01>

F:77-109/Domain: ankyrin repeat homology <AN02>

F:110-142/Domain: ankyrin repeat homology <AN03>

F:143-171/Domain: ankyrin repeat homology <AN04>

F:172-204/Domain: ankyrin repeat homology <AN05>

F:205-237/Domain: ankyrin repeat homology <AN06>

F:238-270/Domain: ankyrin repeat homology <AN07>

F:271-303/Domain: ankyrin repeat homology <AN08>

F:304-336/Domain: ankyrin repeat homology <AN09>

F:337-369/Domain: ankyrin repeat homology <AN10>

F:370-402/Domain: ankyrin repeat homology <AN11>

F:403-435/Domain: ankyrin repeat homology <AN12>

F:436-468/Domain: ankyrin repeat homology <AN13>

F:469-501/Domain: ankyrin repeat homology <AN14>

F:502-534/Domain: ankyrin repeat homology <AN15>

F:535-567/Domain: ankyrin repeat homology <AN16>

F:568-600/Domain: ankyrin repeat homology <AN17>

F:601-633/Domain: ankyrin repeat homology <AN18>

F:634-666/Domain: ankyrin repeat homology <AN19>

F:667-699/Domain: ankyrin repeat homology <AN20>

F:700-732/Domain: ankyrin repeat homology <AN21>

F:733-765/Domain: ankyrin repeat homology <AN22>

F:766-798/Domain: ankyrin repeat homology <AN23>

alignment_scores:

Quality: 290.00

Length: 230

Ratio: 1.959 Gaps: 3
 Percent Similarity: 64.348 Percent Identity: 34.348
 alignment_block:
 US-09-509-775-1 x A35049 ..
 Align seg 1/1 to: A35049 from: 1 to: 1880
 109 GTGTCTTAACCTAATGGTC.....TGCAACCTGGCTACAGCGGAA 149
 ||| : : : : : ||| : : : : :
 432 ValSerAsnValLysValGluThrProLeuHisMetAlaAlaArgAlaGl 448
 150 GCTGGAAGAGTTGAAGGAGAGTATTCTGGCGATAAATCCCTGGCTACTA 199
 ||| : : : : : ||| : : : : :
 448 yHisThrGluValAlaLysTyrLeuLeuGlnAsnLysAlaValAsnA 465
 200 GAACGTACCAGGACAGCAGACTGCATTGCACTGGCGATGCTCAGCTGGA 249
 : : : : : ||| : : : : : ||| : : : : :
 465 lAlaLysAlaLysAspGlnThrProLeuHisCysAlaAlaArgIleGly 481
 250 CATACAGAAATTTGTAATTTTGTTCGAACCTGGAGTGCACGAGTAATGA 299
 ||| : : : : : ||| : : : : : ||| : : : : :
 482 HisThrAsnMetValLysLeuLeuLeuGlnAsnAlaAsnProAsnLe 498
 300 TAAAGACGATGCAGGTGGTCTCTCTTTCATATTGGCGCTTCTCTGGCC 349
 ||| : : : : : ||| : : : : : ||| : : : : :
 498 uAlaThrThrAlaGlyHisThrProLeuHisIleAlaAlaArgGluGlyH 515
 350 GGGATGAGATTGTAAGCCCTTCTGGGAAAAGGTGCTCAAGTGAATGCT 399
 : : : : : ||| : : : : : ||| : : : : :
 515 lValGluThrValLeuAlaLeuLeuGluLysGluAlaSerGlnAlaCys 531
 400 GTCATCAAAATGGCTGCTACTCTCTTACATTTACAGCTTCGAAAAACAG 449
 : : : : : ||| : : : : : ||| : : : : :
 532 MetThrLysLysGlyPheThrProLeuHisValAlaAlaLysTyrGlyL 548
 450 GCATGAGATCGCTGCTCATGTACTTGAAGCGGGGCTAATCCAGATGCTA 499
 : : : : : ||| : : : : : ||| : : : : :
 548 sValArgValAlaGluLeuLeuLeuGluArgAspAlaHisProAsnAlaA 565
 500 AGGACCATTTAGGCTACAGCAATGCACCGGCGAGCAGCAAGGTAAC 549
 ||| : : : : : ||| : : : : : ||| : : : : :
 565 lGlyLysAsnGlyLeuThrProLeuHisValAlaValHisHisAsnAsn 581
 550 TTGAAGATGATTCATATCTCTCTACTACAAAGCATCCACAAACATCCA 599
 ||| : : : : : ||| : : : : : ||| : : : : :
 582 LeuAspIleValLysLeuLeuLeuProArgGlyGlySerProHisSerPr 598
 600 AGACATGAGGTTACACTCTCTTACACTTTAGCCTGTGATGAGGAGAGAG 649
 : : : : : ||| : : : : : ||| : : : : :
 598 oAlaTrpAsnGlyThrProLeuHisIleAlaAlaLysGlnAsnGlnV 615
 650 TGAAGAAGCAAACTGCTGTGTCTCCCAAGGAGCAAGTATTATTCATGAG 699
 ||| : : : : : ||| : : : : : ||| : : : : :
 615 aGluValAlaArgSerLeuLeuGlnThrGlyGlySerAlaAsnAlaGlu 631
 700 AATAAAGAGAAAGACACCCCTGCAAGTGCC.....AAAGTGCC... 741
 : : : : : ||| : : : : : ||| : : : : :
 632 SerValGlnGlyValThrProLeuHisLeuAlaGlnGluGlyHisAl 648
 742CTGGGTTTAACTCAAGAGCAAGTGGTGAAGGT 774
 : : : : : ||| : : : : : ||| : : : : :
 648 aGluMetValAlaLeuLeuSerLysGlnAlaAsnGly 661

seq_name: pir1:SJHUK

seq_documentation_block:

ankyrin 1, erythrocyte splice form 1 - human
 N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
 N:Contains: ankyrin 2.2
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
 C:Accession: S08275; A33219; PC2220; A35443

R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A>Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A:Reference number: S08275; MUID:90158830
A:Accession: S08275
A:Molecule type: mRNA
A:Residues: 1-1881 <LUI>
A:Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A:Accession: A33219
A:Molecule type: protein
A:Residues: 2-7, 'X', 9-17, 'X', 19-20, 'T', 22-30, 733-749, 'A', 751-753; 828-833, 'X', 835-855, 'X'
A>Note: 845-Arg and 1392-Thr were also found
R;Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A>Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrane
A:Reference number: PC2220; MUID:95071348
A:Accession: PC2220
A:Molecule type: protein
A:Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A>Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger
A:Reference number: A35443; MUID:90285190
A:Accession: A35443
A:Molecule type: protein
A:Residues: 'X', 5, 'X', 7-12; 403-417, 'X', 419-422, 'H', 424, 'LQ', 797-800, 'L', 802-814; 862-863;
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; phosphoprotein
F:2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F:2-1512,1675-1881/Product: ankyrin 2, erythrocyte #status predicted <MAT2>
F:2-827/Domain: 89K #status predicted <DOM1>
F:2-827/Region: anion exchange protein binding
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN01>
F:110-142/Domain: ankyrin repeat homology <AN02>
F:143-171/Domain: ankyrin repeat homology <AN03>
F:172-204/Domain: ankyrin repeat homology <AN04>
F:205-237/Domain: ankyrin repeat homology <AN05>
F:238-270/Domain: ankyrin repeat homology <AN06>
F:271-303/Domain: ankyrin repeat homology <AN07>
F:304-336/Domain: ankyrin repeat homology <AN08>
F:337-369/Domain: ankyrin repeat homology <AN09>
F:370-402/Domain: ankyrin repeat homology <AN10>
F:403-435/Domain: ankyrin repeat homology <AN11>
F:436-468/Domain: ankyrin repeat homology <AN12>
F:469-501/Domain: ankyrin repeat homology <AN13>
F:502-534/Domain: ankyrin repeat homology <AN14>
F:535-567/Domain: ankyrin repeat homology <AN15>
F:568-600/Domain: ankyrin repeat homology <AN16>
F:601-633/Domain: ankyrin repeat homology <AN17>
F:634-666/Domain: ankyrin repeat homology <AN18>
F:667-699/Domain: ankyrin repeat homology <AN19>
F:700-732/Domain: ankyrin repeat homology <AN20>
F:733-765/Domain: ankyrin repeat homology <AN21>
F:766-798/Domain: ankyrin repeat homology <AN22>
F:828-1382/Domain: ankyrin repeat homology <AN23>
F:828-1382/Region: 62K #status predicted <DOM2>
F:1383-1881/Domain: 55K #status predicted <DOM3>

alignment_scores:	
Quality:	290.00
Ratio:	1.959
Percent Similarity:	64.348
Length:	230
Gaps:	3
Percent Identity:	34.348

alignment_block:

US-09-509-775-1 x SJHUK

```

seq_name: pir2:A57282
seq_documentation_block:
  C:Species: Caenorhabditis elegans (fragment)
  C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 20
  C:Accession: A57282
  R:Otsuka, A.J.; Franco, R.; Yang, B.; Shlm, K.H.; Tang, L.Z.; Zhang,
  J. Cell Biol. 129, 1081-1092, 1995
  A:title: An ankyrin-related gene (unc-44) is necessary for proper
  A:Reference number: A57282; MUID:95263663
  A:Accession: A57282
  A>Status: Preliminary
  A:Molecule type: DNA
  A:Residues: 1-1786 <OTS>

```

A:Cross-references: GB:U21734; NID:G790607; PIDN:AAA85854.1; PID:G790608
 C:Genetics:
 A:Gene: unc-44
 A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1725/1
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
 F:164-182/Domain: ankyrin repeat homology <AN04>
 F:358-390/Domain: ankyrin repeat homology <AN1>
 F:391-423/Domain: ankyrin repeat homology <AN11>

alignment_scores:
 Quality: 288.50 Length: 211
 Ratio: 2.003 Gaps: 4
 Percent Similarity: 68.246 Percent Identity: 36.019
 alignment_block:
 US-09-509-775-1 x A57282 ..
 Align seg 1/1 to: A57282 from: 1 to: 1786

```

154 GAAGAGTTCAGGAGAGATATTCTG.....GCCGATAAATCCCTGCG 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
537 GluGluValAlaGlyLeuLeuLeuAspHisAsnAlaAspLysThrLeuLe 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 TACTAGAACTGACGAGGAGAGAACTGCATTGCCTGGGCTGCTCAG 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 uThrLys.....LysGlyPheThrProLeuHisLeuAlaSerLysT 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
245 CTGGACATACAGAAATGTTGAATTTTGTTCGCAACTTGGAGTGCAGTG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 yrGlyAsnLeuGluValValArgLeuLeuLeuGluArgGlyThrProVal 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
295 AATGATAAGACGATCCAGGTTGGTCTCTCTCATATTGCGGCTCTCGC 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
584 AspIleGluGlyLysAsnGlnValThrProLeuHisValAlaAlaHisTy 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
345 TGCCCGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTGA 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
600 rAsnAsnAspLysValAlaMetLeuLeuLeuGluAsnGlyAlaSerAlaL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
395 ATGCTGTCAATCAAAATGCTGTACCTTTCCTTACATATGCGAGTTCGAAA 444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
617 ysAlaAlaAlaLysAsnGlyTyThrProLeuHisIleAlaAlaLysLys 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 AACAGCATGAGATCGCTCATGTTCTGCTTACTGGAAGCGGGCTTAATCCAGA 494
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
634 AsnGlnMetGluIleAlaSerThrLeuLeuGlnPheLysAlaAspProAs 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
495 TGCTAAGGACCATTTATGAGGCTACAGCAATGCGCGGAGCGCAAGG 544
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
650 nAlaLysSerArgAlaGlyPheThrProLeuHisLeuSerAlaGlnGluG 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
545 GTAACCTGAAGATGATTATATCCTTCTTCTTACTACAAAGCATCCACAAC 594
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
667 lYHisLysGluIleSerGlyLeuLeuIleGluAsnGlySerAspValGly 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
595 ATCCAGACACTGAGGTAAACACTCTCTTACACTTACCTGTGATCCAGA 644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
684 AlaLysAlaAsnAsnGlyLeuThrAlaMetHisLeuCysAlaGlnGluAs 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
645 GAGAGTGGGAAGCAAGCAAACTGCTGTGTGTCGCCAAGGAGCAAGTAT 694
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
700 pHIsValProValAlaGlnIleLeuTyrrAsnAsnGlyAlaGluIleAsnS 717
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
695 TTGAGATAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 738
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
717 erLysThrAsnAlaGlyTyThrProLeuHisValAlaCysHisPheGly 733
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
739 GCCTGGGTTTAACTCACTCAAGAGATGGTGAA 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
734 GlnLeuAsnMetVal...LysPheLeuValGlu 743
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: p1r2:T15346

seq_documentation_block:
 elegans ankyrin-related unc-44 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 C:Accession: T15346
 R:Gattung, S.
 submitted to the EMBL Data Library, February 1996
 A:Description: The sequence of C. elegans cosmid B0350.
 A:Reference number: Z18332
 A:Accession: T15346
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1815 <GAT>
 A:Cross-references: EMBL:U50071; NID:gl208871; PID:gl208875; PIDN:AAA93445.1; CESP:un
 C:Genetics:
 A:Gene: CESP:unc-44
 A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
 C:Superfamily: ankyrin; ankyrin repeat homology
 F:358-390/Domain: ankyrin repeat homology <ANR>

alignment_scores:
 Quality: 288.50 Length: 211
 Ratio: 2.003 Gaps: 4
 Percent Similarity: 68.246 Percent Identity: 36.019
 alignment_block:
 US-09-509-775-1 x T15346 ..
 Align seg 1/1 to: T15346 from: 1 to: 1815

```

154 GAAGAGTTCAGGAGAGATATTCTG.....GCCGATAAATCCCTGCG 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
537 GluGluValAlaGlyLeuLeuLeuAspHisAsnAlaAspLysThrLeuLe 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 TACTAGAACTGACGAGGAGAGAACTGCATTGCCTGGGCTGCTCAG 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 uThrLys.....LysGlyPheThrProLeuHisLeuAlaSerLysT 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
245 CTGGACATACAGAAATGTTGAATTTTGTTCGCAACTTGGAGTGCAGTG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 yrGlyAsnLeuGluValValArgLeuLeuLeuGluArgGlyThrProVal 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
295 AATGATAAGACGATCCAGGTTGGTCTCTCTTACTTATGCGGCTCTCGC 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
584 AspIleGluGlyLysAsnGlnValThrProLeuHisValAlaAlaHisTy 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
345 TGCCCGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTGA 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
600 rAsnAsnAspLysValAlaMetLeuLeuLeuGluAsnGlyAlaSerAlaL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
395 ATGCTGTCAATCAAAATGCTGTACCTTTCCTTACATATGCGAGTTCGAAA 444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
617 ysAlaAlaAlaLysAsnGlyTyThrProLeuHisIleAlaAlaLysLys 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 AACAGCATGAGATCGCTCATGTTCTGCTTACTGGAAGCGGGCTTAATCCAGA 494
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
634 AsnGlnMetGluIleAlaSerThrLeuLeuGlnPheLysAlaAspProAs 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
495 TGCTAAGGACCATTTATGAGGCTACAGCAATGCGCGGAGCGCAAGG 544
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
650 nAlaLysSerArgAlaGlyPheThrProLeuHisLeuSerAlaGlnGluG 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
545 GTAACCTGAAGATGATTATATCCTTCTTCTTACTACAAAGCATCCACAAC 594
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
667 lYHisLysGluIleSerGlyLeuLeuIleGluAsnGlySerAspValGly 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
595 ATCCAGACACTGAGGTAAACACTCTCTTACACTTACCTGTGATCCAGA 644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
684 AlaLysAlaAsnAsnGlyLeuThrAlaMetHisLeuCysAlaGlnGluAs 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
645 GAGAGTGGGAAGCAAGCAAACTGCTGTGTGTCGCCAAGGAGCAAGTAT 694
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


OM of: US-09-509-775-1 to: SwissProt_39.* out_format : pfs
Date: Aug 13, 2001 8:37 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+np.model -DRV=xlp
-Q/cgn2_1/USPTO_509509775/runat_13082001_074528_13094/app_query.fasta_1.848
-DB=SwissProt_39 -QFMT=fastan -SUFFIX=rs -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QCAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=200000000
-USER=US09509775_ECNG1_1_28 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-509-775-1
Query length: 780
Database: SwissProt_39.*
Database sequences: 33435
Database length: 3425486
Search time (sec): 62.940000

score_list:	Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SwissProt_39:ANK2_HUMAN	+	294.00	456.18	2.3e-19	3924	! Q01484 homo sapiens (human)
SwissProt_39:Y64X_YEAST	+	291.00	475.54	3.1e-19	228	! P50086 saccharomyces cerevisiae
SwissProt_39:ANK1_HUMAN	+	290.00	455.83	5.1e-19	1880	! P16157 homo sapiens (human)
SwissProt_39:ANK1_MOUSE	+	287.50	451.76	8.1e-19	1862	! Q02357 mus musculus (mouse)
SwissProt_39:YD57_SCHPO	+	273.50	446.24	1.4e-17	234	! Q10311 schizosaccharomyces
SwissProt_39:PA26_MOUSE	+	254.50	404.68	9.0e-16	732	! P57819 mus musculus (mouse)
SwissProt_39:ANK3_HUMAN	+	248.50	393.85	3.2e-15	832	! P57078 homo sapiens (human)
SwissProt_39:PA26_RAT	+	247.50	393.06	4.0e-15	751	! P97570 rattus norvegicus (rat)
SwissProt_39:LATA_LATMA	+	247.50	387.73	4.2e-15	1401	! P23631 latrodectus mactans
SwissProt_39:DAPK_HUMAN	+	247.00	386.71	4.7e-15	1431	! P53355 homo sapiens (human)
SwissProt_39:PA26_HUMAN	+	237.50	375.84	3.4e-14	806	! Q06733 homo sapiens (human)
SwissProt_39:AKR_ATHAT	+	236.50	379.38	3.9e-14	439	! Q05753 arabidopsis thaliana
SwissProt_39:ANKH_CHRVI	+	225.00	362.90	4.4e-13	323	! Q06527 chromatium vinosum
SwissProt_39:KBF2_HUMAN	+	224.50	353.32	5.4e-13	898	! Q00653 homo sapiens (human)
SwissProt_39:ILK_MOUSE	+	217.50	347.56	2.3e-12	452	! O55222 mus musculus (mouse)
SwissProt_39:PEM1_CAEEL	+	215.00	340.22	4.0e-12	656	! P17221 caenorhabditis elegans
SwissProt_39:NOTC_DROME	+	215.00	328.10	4.6e-12	2703	! P07207 drosophila melanogaster
SwissProt_39:ITK3_MOUSE	+	214.00	327.76	5.6e-12	2318	! Q61982 mus musculus (mouse)
SwissProt_39:ILK_CAVPO	+	212.50	339.27	6.6e-12	451	! P57044 cavia porcellus (guinea pig)
SwissProt_39:NTCL_HUMAN	+	212.00	323.68	8.7e-12	2531	! Q07008 rattus norvegicus (rat)
SwissProt_39:ILK1_HUMAN	+	211.50	337.59	8.1e-12	452	! Q13418 homo sapiens (human)
SwissProt_39:ILK2_HUMAN	+	211.50	337.59	8.1e-12	452	! P57043 homo sapiens (human)
SwissProt_39:NTC1_HUMAN	+	209.50	319.83	1.5e-11	2444	! P46531 homo sapiens (human)
SwissProt_39:NTC1_MOUSE	+	208.50	320.03	1.8e-11	1964	! P31695 mus musculus (mouse)
SwissProt_39:NOTC_BRARE	+	207.50	316.53	2.2e-11	2437	! P46530 brachydanio rerio (zebrafish)
SwissProt_39:RN5A_MOUSE	+	202.00	318.32	6.4e-11	679	! Q05921 mus musculus (mouse)
SwissProt_39:BC13_HUMAN	+	201.00	320.26	7.1e-11	446	! P20749 homo sapiens (human)
SwissProt_39:NOTC_XENLA	+	200.00	303.76	1.1e-10	2524	! P21783 xenopus laevis (afrotoad)
SwissProt_39:KBF1_CHICK	+	198.00	308.50	1.6e-10	984	! Q04861 gallus gallus (chicken)
SwissProt_39:RN5A_HUMAN	+	194.00	304.28	3.6e-10	741	! Q05823 homo sapiens (human)
SwissProt_39:AKR1_YEAST	+	194.00	304.02	3.6e-10	764	! P39010 saccharomyces cerevisiae
SwissProt_39:GABB_HUMAN	+	189.50	303.30	8.6e-10	347	! Q06545 homo sapiens (human)
SwissProt_39:GABB_MOUSE	+	189.50	302.45	8.7e-10	383	! Q06547 homo sapiens (human)
SwissProt_39:KBF1_HUMAN	+	188.00	292.02	1.3e-09	969	! P19838 homo sapiens (human)
SwissProt_39:CACT_DROME	+	187.50	296.85	1.4e-09	500	! Q03017 drosophila melanogaster
SwissProt_39:KBF1_MOUSE	+	187.50	291.17	1.5e-09	971	! P25799 mus musculus (mouse)
SwissProt_39:L112_CAEEL	+	187.00	287.03	1.7e-09	1429	! P14585 caenorhabditis elegans
SwissProt_39:GABB_MOUSE	+	186.50	297.49	1.6e-09	382	! Q00420 mus musculus (mouse)
SwissProt_39:KBF1_RAT	+	186.00	293.99	1.9e-09	522	! Q63369 rattus norvegicus (rat)

SwissProt_39:NTCL_MOUSE + 183.50 276.33 3.8e-09 2531 ! Q01705 mus musculus (mouse)
SwissProt_39:CDN6_HUMAN + 183.00 298.71 3.2e-09 168 ! P42773 homo sapiens (human)
SwissProt_39:HT16_HYDAT + 182.50 284.99 4.1e-09 757 ! P53356 hydra attenuata (hydra)
SwissProt_39:YA2A_SCHPO + 181.50 284.74 5.0e-09 642 ! Q09701 schizosaccharomyces
SwissProt_39:RFKX_HUMAN + 181.00 291.65 5.1e-09 260 ! O14593 homo sapiens (human)

seq_name: SwissProt_39:ANK2_HUMAN

seq_documentation_block:

AD ANK2_HUMAN STANDARD; PRT; 3924 AA.
ID Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeh E., Bennett V.;
RT "440-KD ankyrinB: structure of the major developmentally regulated domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [3]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E., Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -1- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL ELEMENTS: THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOSOLIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL CELLS THROUGHOUT THE BRAIN.
CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE AND FUNCTION (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; Z26634; CAB42644.1; -
DR EMBL; X56957; CAA40278.1; -
DR EMBL; X56958; CAA40279.2; -
DR EMBL; M37123; AAG62828.1; -
DR PIR; S14533; S14533
DR PIR; A39643; A39643
DR PIR; B39643; B39643
DR PIR; S14569; S14569
DR HSSP; Q00420; LANC
DR MIM; 106410; -
DR InterPro; IPR000488; -
DR InterPro; IPR000906; -
DR InterPro; IPR002110; -
DR Pfam; PF00791; Z05; 1
DR Pfam; PF00023; ank; 22
DR Pfam; PF00531; death; 1
DR PROSITE; PS50088; ANK_REPEAT; 20
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1
DR PROSITE; PS50017; DEATH_DOMAIN; 1
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation; Multigene family
FT REPEAT 63 92
FT REPEAT 96 125 ANK 2
FT REPEAT 129 158 ANK 3
FT REPEAT 162 191 ANK 4
FT REPEAT 193 220 ANK 5
FT REPEAT 232 261 ANK 6
FT REPEAT 265 294 ANK 7
FT REPEAT 298 327 ANK 8
FT REPEAT 331 360 ANK 9
FT REPEAT 364 393 ANK 10
FT REPEAT 397 426 ANK 11
FT REPEAT 430 459 ANK 12
FT REPEAT 463 492 ANK 13
FT REPEAT 496 525 ANK 14
FT REPEAT 529 558 ANK 15
FT REPEAT 562 591 ANK 16
FT REPEAT 595 624 ANK 17
FT REPEAT 628 657 ANK 18
FT REPEAT 661 690 ANK 19
FT REPEAT 694 723 ANK 20
FT REPEAT 727 756 ANK 21
FT REPEAT 760 789 ANK 22
FT REPEAT 793 822 ANK 23
FT DOMAIN 1773 1950 REPEAT-RICH REGION.
FT REPEAT 1773 1784 REPEAT A.
FT REPEAT 1785 1796 REPEAT A.
FT REPEAT 1797 1808 REPEAT A.
FT REPEAT 1809 1820 REPEAT A.
FT REPEAT 1821 1832 REPEAT A.
FT REPEAT 1833 1844 REPEAT A.
FT REPEAT 1845 1856 REPEAT A.
FT REPEAT 1857 1867 REPEAT A. (APPROXIMATE).
FT REPEAT 1868 1879 REPEAT A.
FT REPEAT 1880 1891 REPEAT A.
FT REPEAT 1892 1902 REPEAT A. (APPROXIMATE).
FT REPEAT 1903 1914 REPEAT A.
FT REPEAT 1915 1926 REPEAT A.
FT REPEAT 1927 1938 REPEAT A.
FT REPEAT 1939 1950 REPEAT A.
FT DOMAIN 3536 3620 DEATH DOMAIN.
FT VARSPPLIC 1039 1039 (IN ISOFORM 1).
FT VARSPPLIC 1444 3528 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 475 476 GO -> PE (IN REF. 4).
FT CONFLICT 971 971 I -> S (IN REF. 2).
FT CONFLICT 3581 3582 QY -> HA (IN REF. 2).
FT CONFLICT 3586 3586 I -> Y (IN REF. 2).
FT SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;
```

alignment_scores:

Quality: 294.00

Length: 232

```
Ratio: 2.146 Gaps: 2
Percent Similarity: 59.052 Percent Identity: 31.897
alignment_block:
US-09-509-775-1 x ANK2_HUMAN
Align seg 1/1 to: ANK2_HUMAN from: 1 to: 3924
136 GCCTACAGCGGGGAAGCTGGAAGAGTTCAGAGGAGTAGTATTGCGCCGATAA 185
||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
472 AlaArgAlaGlyGlnValValArg...CysLeuLeuArgAsnG1 487
186 ATCCCTGGCTACTAGAACTGACAGGACGACGACAGCAATGCTTCACATGGG 235
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
487 yAlaLeuValAspAlaArgAlaArgGluGluGlnThrProLeuHisIleA 504
236 CATGCTCAGCTGCACATACAGAAATGTTCAATTTTGTGCAA..... 279
||| :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
504 laSerArgLeuGlyLysThrGluIleValGlnLeuLeuGlnHisMet 520
279 ..... 279
521 AlaHisProAspAlaAlaThrThrAsnGlyTyrThrProLeuHisIleSe 537
280 .....CTTGGAG 286
|||:
537 rAlaArgGluGlyGlnValAspValAlaSerValLeuLeuGluAlaGlyA 554
287 TGCCAGTGAATGATAAAGAGCATGCAGGTGGTCTCTCTTCATATTGCG 336
: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
554 laAlaHisSerLeuAlaThrLysLysGlyPheThrProLeuHisValAla 570
337 GCTCTGCTGGCGGGATGAGATTGTAAGAGCCCTTCTGGGAAAGGTGC 386
|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
571 AlaLysTyrGlySerLeuAspValAlaLysLeuLeuGlnArgArgAl 587
387 TCAAGTGAATGCTGCAATCAAAATGGCTGTACTCCCTTACATATTGCG 436
: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
587 aAlaAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAla 604
437 CTTGCAAAACAGCATGAGATCGCTGCTCATGTTACTTGGGAGCGGGCT 486
|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
604 laHisTyrAspAsnGlnLysValAlaLeuLeuLeuGluLysGlyAla 620
487 AATCCAGATGCTAAGGACCATATTAGGCTACAGCAATGCACCGGCGAGC 536
:|||||: ||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
621 SerProHisAlaThrAlaLysAsnGlyTyrThrProLeuHisIleAla 637
537 AGCCAGGGTAACTTGAAGATGATTTCATATCTCTGTACTACAAAGCAT 586
: |||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
637 aLysLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnTyrGlyAla 654
587 CCACAACATCCAGACACGTAGGGAAGCACTCTCTACACTTAGCGTGT 636
:|||||: ||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
654 luThrAsnIleValThrLysGlnGlyValThrProLeuHisLeuAlaSer 670
637 GATCAGAGAGAGTGGGAAGCAAAACTGCTGTGTCTCCCAAGGAGCAAG 686
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
671 GlnGluGlyHisThrAspMetValThrLeuLeuAspLysGlyAlaAs 687
687 TATTTCATTGAGAAATAAGAAAGAACACCCCTGCAAGTGGCC 732
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
687 nIleHisMetSerThrLysSerGlyLeuThrSerLeuHisLeuAla 702
seq_name: SwissProt_39:YG4X_YEAST
seq_documentation_block:
ID YG4X_YEAST STANDARD; PRT; 228 AA.
AC P50086;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 25.6 KDA PROTEIN IN SM11-PHO81 INTERGENIC REGION.
```

GN YGR232W OR G8564.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96267763; PubMed=8701610;
RA van der Aart O.J.M., Kleine K., Steensma H.Y.;
RT "Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM1-PHO81-YHB4-
RT PFK1 region from the right arm of Saccharomyces cerevisiae chromosome
RT VII.";
RL Yeast 12:385-390(1996).
CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; X87941; CAA61182.1; -;
DR EMBL; Z73017; CAA97260.1; -;
DR SGD; S0003464; YGR232W.
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; ANK repeat; Repeat.
FT REPEAT 1 30 ANK 1.
FT REPEAT 35 64 ANK 2.
FT REPEAT 71 100 ANK 3.
FT REPEAT 106 135 ANK 4.
FT REPEAT 139 168 ANK 5.
FT REPEAT 173 203 ANK 6.
SQ SEQUENCE 228 AA; 25616 MW; 6B30C0DD034A6289 CRC64;

alignment_scores:
Quality: 291.00 Length: 210
Ratio: 2.109 Gaps: 4
Percent Similarity: 65.714 Percent Identity: 35.714

alignment_block:

US-09-509-775-1 x YG4X_YEAST

Align seg 1/1 to: YG4X_YEAST from: 1 to: 228

109 GTGCTTAACCTAATGGTCTGCAACCTGCCTACAGCGGAGCTGGAGA 158
1 MetSerAsnTyrProLeuHisGlnAlaCysMetGluAsnGluPheLeu 17
159 GTTCAAGCAGAGTAGTTCTGGCCGATAAATCCCTACTAGACTGACC 208
17 svaGlnGluLeuLeuHisSerLysProSerLeuLeuGlnLysAspG 34
209 AGGACAGCAGCAACTGCGATTTGCACCTGGCGATCCTCAGCTGGACATACAGAA 258
34 InAspGlyArgIleProLeuHisTrpSerValSerPheGlnAlaHisGlu 50
259 ATTCTGAATTTTGTTCACAACTGGAGTGCAGTGAAT.....GA 299
51 IleThrSerPheLeuSerLysMetGluAsnValAlaSerPheGlnAlaHisGly 67
300 TAAACAGATCCAGCTGGTCTCTCTTCATATTGCGGCTCTCTGCTGCC 349
67 rProAspAspSerGlyTrpThrProPheHisIleAlaCysSerValGly 84
350 GGGATGAGATTGTAAGACCTCTTGGGAAAGGTGCTCAA.....GTG 393

84 snLeuGluValValLysSerLeuTyrAspArgProLeuLysProAspLeu 100
394 AATGCTCTCAATCAAAATGGCTGACTCCCTACATTATGAGCTTGA 443
101 AsnLysIleThrAsnGlnGlyValThrCysLeuHisLeuAlaValGly 117
444 AACAGCAGTCAGATCGCTGCTCATCTTACTGGAAGCGGGCTAATCCAG 493
117 sLysTrpPheGluValSerGlnPheLeuIleGluAsnGlyAlaSerVal 134
494 ATGCTAAGGACCATATGAGCTACACATGACCGGCGAGCAGCCAAAG 543
134 rgileLysAspLysPheAsnGlnIleProLeuHisArgAlaSerVal 150
544 GGTAACTTGAAGATGATTCATATCTCTCTG...TACTACAAAGCATCCAC 590
151 GlySerLeuLysLeuIleGluLeuLeuCysGlyLeuGlyLysSerAlaVal 167
591 AAACATCAAGACACTGAGGCTAACACTCTCTACACTTACCTGCTGATG 640
167 lasnTrpGlnAspLysGlnGlyTrpThrProLeuPheHisAlaLeuAlaG 184
641 AGGAGAGAGTGGGAAGACCAAACTGCTGTGTCCTCCAGGAGCAAGTATT 690
184 luGlyHisGlyAspAlaValLeuLeuValGluLysTyrGlyAlaGlu 200
691 TAC.....ATTGAGAATAAAGAGAAAG 714
201 TyrAspLeuValAspAsnLysGlyAlaLys 210
seq_name: SwissProt_39:ANK1_HUMAN
seq_documentation_block:
ID ANK1_HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins.";
RL Nature 344:36-42(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175370; PubMed=1689849;
RA Lambert S., Yu H., Pichai J.T., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RP VARIANT HS ILE-462.
RX MEDLINE=96225450; PubMed=8640229;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis.";
RL Nat. Genet. 13:214-218(1996).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.

CC	ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC	CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC	THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC	PLASMA MEMBRANE.
CC	-1- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF ANKYRIN ARE PRODUCED BY
CC	ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ANKYRIN
CC	VARIANT 2.1.
CC	-1- PTM: REGULATED BY PHOSPHORYLATION.
CC	-1- PTM: ACYLATED BY PALMITIC ACID GROUP(S).
CC	-1- DISEASE: DEFECTS IN ANK1 ARE THE CAUSE OF DOMINANT AND RECESSIVE
CC	HEREDITARY SPHEROCYTOSIS (HS).
CC	-1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to licensee@isb-sib.ch).
CC	-----
DR	EMBL; X16609; CAA34610.1; -
DR	ENBL; M28880; AAA51732.1; -
DR	PIR; S08275; SJHUK.
DR	PIR; A35049; A35049.
DR	HSSP; Q00420; IAWC.
DR	MIM; 182900; -
DR	InterPro; IPR000488; -
DR	InterPro; IPR000906; -
DR	InterPro; IPR002110; -
DR	Pfam; PF00791; ZU5; 1.
DR	Pfam; PF00023; ank; 22.
DR	Pfam; PF00531; death; 1.
DR	PROSITE; PS50088; ANK_REPEAT; 20.
DR	PROSITE; PS50297; ANK_REP_REGION; 1.
DR	PROSITE; PS50017; DEATH_DOMAIN; 1.
KW	Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW	Phosphorylation; Lipoprotein; Multigene family; Disease mutation;
KW	Elliptycotosis; Polymorphism.
FT	INIT_MET 0 0
FT	DOMAIN 1 826
FT	89 KDA DOMAIN (ANION EXCHANGE PROTEIN
FT	BINDING DOMAIN).
FT	DOMAIN 827 1381
FT	62 KDA DOMAIN (SPECTRIN BINDING
FT	DOMAIN).
FT	DOMAIN 1382 1880
FT	55 KDA REGULATORY DOMAIN (REGULATES
FT	THE BINDING OF ANKYRIN TO SPECTRIN
FT	AND THE BAND 3 PROTEIN).
FT	ANK 1.
FT	REPEAT 43 72
FT	REPEAT 76 105
FT	ANK 2.
FT	REPEAT 109 138
FT	ANK 3.
FT	REPEAT 142 171
FT	ANK 4.
FT	REPEAT 173 200
FT	ANK 5.
FT	REPEAT 204 233
FT	ANK 6.
FT	REPEAT 237 266
FT	ANK 7.
FT	REPEAT 270 299
FT	ANK 8.
FT	REPEAT 303 332
FT	ANK 9.
FT	REPEAT 369 398
FT	ANK 10.
FT	REPEAT 402 431
FT	ANK 11.
FT	REPEAT 435 464
FT	ANK 12.
FT	REPEAT 468 497
FT	ANK 13.
FT	REPEAT 501 530
FT	ANK 14.
FT	REPEAT 534 563
FT	ANK 15.
FT	REPEAT 567 596
FT	ANK 16.
FT	REPEAT 600 629
FT	ANK 17.
FT	REPEAT 633 662
FT	ANK 18.
FT	REPEAT 666 695
FT	ANK 19.
FT	REPEAT 699 728
FT	ANK 20.
FT	REPEAT 732 761
FT	ANK 21.
FT	REPEAT 765 794
FT	ANK 22.
FT	REPEAT 1402 1486
FT	DEATH DOMAIN.
FT	DOMAIN 1512 1873
FT	MISSING (IN ISOFORM 2.2).
FT	VARSPLIC


```
CC HEART AND SKELETAL MUSCLE.
CC -1- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U51898; AAC53136.1;
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hydrolyase; Lipid degradation; Repeat; ANK repeat.
FT REPEAT 150 180
FT REPEAT 184 214 ANK 1.
FT REPEAT 218 247 ANK 2.
FT REPEAT 250 280 ANK 3.
FT REPEAT 285 311 ANK 4.
FT REPEAT 315 344 ANK 5.
FT REPEAT 348 377 ANK 6.
FT ACT_SITE 464 464 ANK 7.
FT ACT_SITE 464 464 POTENTIAL.
SQ SEQUENCE 751 AA; 83582 MW; 393BBBADA7FCC99B CRC64;

alignment_scores:
  Quality: 247.50      Length: 246
  Ratio: 1.695        Gaps: 6
  Percent Similarity: 59.350      Percent Identity: 28.455

alignment_block:
US-09-509-775-1 x PA26_RAT ..
Align seg 1/1 to: PA26_RAT from: 1 to: 751

100 GAGGGGTGTGTCTTAACCTAATGCTGCTCAACCTGGCTGCTACAGCGGAA 149
150 GluGlyCysThrProLeuHisLeuAlaCysArgLysGlyAspSerGlu1 166
150 GCTGGAAGAGTGTGAAGCAGAGTATTCTGCGGATAAATCCCTGGCTACTA 199
166 eleuValGluLeuValGlnTyrCysHisAlaGlnMetaspVal..... 180
200 GAACCTGACGAGCAGCAGACAGTGCATGCTGCTGCTGCTGCTGCTGCTGGA 249
181 ...ThrAspAsnLysGlyGluThrAlaPheHisTyrAlaValGlnGlyAsp 196
250 CATACAGAAATGTTGAATTTTG...TTGCCAACTTGGAGTGCCAGTGAA 296
197 AsnProGlnValLeuGlnLeuLeuLysAsnAlaSerAlaGlyLeuAs 213
297 TGATAAAGCAGTGCAGGTGGTCTCTCTTCATATTGCGGCTTCGCTG 346
213 nGlnValAsnAsnGlnGlyLeuThrProLeuHisLeuAlaCysGlnMetG 230
347 GCGGGATGAGATGTGAAGCCCTTCTGGGAAAGGTGCTCAAGTGAAT 396
230 lLysGlnGluMetValArgValLeuLeuLeuCysAsnAlaArgCysAsn 246
397 GCTGTC.....AATCA 407
247 IleMetGlyProGlyPhePheProIleHisThrAlaMetLysPheSerG 263
408 AATAGCTGT..... 417
263 nLysGlyCysAlaGluMetIleIleSerMetAspSerAsnGlnIleHis 280
418 .....ACTCCCTTACATTATGCAAGTCCGAA 444
:::|||||:::|||||
```

```
280 erLysAspProArgTyrGlyAlaSerProLeuHisTrpAla.....Lys 294
445 ACAGGCGATGATCGCTCTCATGTTACTTGGAGCGGCGCTAATCCAGA 494
295 AsnAla...GluMetAlaArgMetLeuLeuLysArgGlyCysAspValAs 310
495 TGCTAAGGACCATTTATGAGCGCTACAGCAATGCACCGGCGAGCAAGG 544
310 pSerThrSerAlaSerGlyAsnThrAlaLeuHisValAlaValThrArg 327
545 GTAACCTGAAGATGATTCATATCTCTTCTACTACAAAGCATCCACAAC 594
327 snArgPheAspCysValMetValLeuLeuThrTyrGlyAlaAsnAlaGly 343
595 ATCCAGACACTGAGGTAACACTCTCTACACTTAGCTGCTGATGATGAG 644
344 AlaArgGlyGluHisGlyAsnThrProLeuHisLeuAlaMetSerLysAs 360
645 GAGAGTGGAAGAGCAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTACA 694
360 pAsnMetGluMetValLysAlaLeuIleValPheGlyAlaGluValAsp 377
695 TTGAGAATAAGAAAGAAAGACACCCCTGCAAGTGCC 732
377 hrProAsnAspPheGlyGluThrProAlaPheIleAla 389
seq_name: SwissProt_39:LATA_LATMA

seq_documentation_block:
ID LATA_LATMA STANDARD; PRT; 1401 AA.
AC P23631;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ALPHA-LATROTOXIN PRECURSOR.
OS Latrodectus mactans (Black widow spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Therididae; Latrodectus.
OX NCBI_TaxID=6924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRDECIMGUTATUS; TISSUE=Venom;
RX MEDLINE=91031994; PubMed=1977615;
RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black
RT widow spider venom.";
RL FEBS Lett. 270:127-131(1990).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=91362695; PubMed=1888339;
RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT "Structure of tryptic fragments of a neurotoxin from black widow
RT spider venom.";
RL Bioorg. Khim. 17:437-441(1991).
CC -1- FUNCTION: CAUSES SYNAPTIC VESICLES EXOCYTOSIS AND
CC NEUROTRANSMITTER RELEASE FROM PRESYNAPTIC NERVE TERMINALS.
CC -1- SIMILARITY: CONTAINS 21 ANK REPEATS.
DR PIR; S11527; S11527.
DR HSSP; Q00420; LAMC.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 17.
DR PROSITE; PS50088; ANK_REPEAT; 11.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Toxin; Neurotoxin; Repeat; ANK repeat; Venom; Signal.
FT SIGNAL 1 20
FT CHAIN 21 1401 ALPHA-LATROTOXIN.
FT REPEAT 490 521 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 559 589 ANK 3.
FT REPEAT 593 622 ANK 4.
FT REPEAT 626 656 ANK 5.
FT REPEAT 660 690 ANK 6.
```

```

FT REPEAT      695      723      ANK 7.
FT REPEAT      729      758      ANK 8.
FT REPEAT      762      791      ANK 9.
FT REPEAT      795      824      ANK 10.
FT REPEAT      828      857      ANK 11.
FT REPEAT      862      891      ANK 12.
FT REPEAT      895      924      ANK 13.
FT REPEAT      928      957      ANK 14.
FT REPEAT      971      1003      ANK 15.
FT REPEAT      1003      1033      ANK 16.
FT REPEAT      1035      1064      ANK 17.
FT REPEAT      1068      1097      ANK 18.
FT REPEAT      1101      1131      ANK 19.
FT REPEAT      1137      1166      ANK 20.
FT REPEAT      1170      1199      ANK 21.
SQ SEQUENCE    1401 AA; 156834 MW; 21CE198B80172568 CRC64;

```

```

alignment_scores:
  Quality: 247.50      Length: 181
  Ratio: 1.980         Gaps: 2
  Percent Similarity: 69.061      Percent Identity: 31.492

```

alignment_block:

US-09-509-775-1 x LATA_LATMA

Align seg 1/1 to: LATA_LATMA from: 1 to: 1401

```

220 ACTGATGTCAGTGGGCTGCTCAGTCGACATACAGAAATGTGTAAT 269
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
732 ThrProLeuHisLeuAlaValIleGlnGlyGlySerGlnIleLeuSerLe 748

270 TTTGTGCAACTGTGAGTCCAGTGAATGATAAAGACGATCAGCTTGGT 319
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 uMetPheAspIleGlyValAsnIleGluGlnLysThrAspGluLysTyrT 765

320 CTCCTCTCATATGCGGCTTCTGCTGCCGGGATGAGATTGTAAGGCC 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
765 hrProLeuHisLeuAlaMetSerLysTyrProGluLeuIleGlnIle 781

370 CTCCTGGGAAAGGTCFCAAGTGAATGCTCTCAATCAAAATGGCTGAC 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
782 LeuLeuAspGlnGlySerAsnPheGluAlaLysThrAsnSerGlyAlaTh 798

420 TCCCTTACATATTACGCTTGTGAAACAGCAGCATGCTGCTGTCATGT 469
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
798 rProLeuHisLeuAlaThrPheLysGlyLysSerGlnAlaLeuIleL 815

470 TACTGGAAGCGGGCTAATCCAGATGCTAAGGACCATATGAGGCTACA 519
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
815 euLeuAsnAsnGluValAsnTrpArgAspThrAspGluAsnGlyGlnMet 831

520 GCAATGTCACCGGCGGAGCCAGCAAGGTAACTGAAGATGATTCATATCT 569
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
832 ProfileHisGlyAlaAlaMetThrGlyLeuLeuAspValAlaGlnAlaI 848

570 TCTGTACTACAAAGCATCC...ACAACATCCACAGACTGAGGGTACAC 616
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
848 eileSerileAspAlaThrValValAspIleGluAspLysAsnSerAspT 865

617 CTCCTCTACACTACCTGTGATGAGGAGAGTGGAGAGCAAGAACTG 666
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
865 hrProLeuAsnLeuAlaAlaGlnAsnSerHisIleAspValIleLysTyr 881

667 CTGTGTCTCCAGGACGAGTATTACATTGAGATAAAGAGAAAGAC 716
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
882 PheIleAspGlnGlyAlaAspIleAsnThrArgAsnLysLysGlyLeuAl 898

717 ACCCTGTCAAGTG.....GCCAAGGTGGCTGGTTTAAAT 753
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
898 aProLeuLeuAlaPheSerLysLysGlyAsnLeuAspMetVal 912

```

seq_name: SwissProt_39:DAPK_HUMAN

seq_documentation_block:

```

ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.1-) (DAP KINASE 1).
GN DAPK1 OR DAPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129831; PubMed=7828849;
RA Weiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD
RT protein as potential mediators of the gamma interferon-induced cell
RT death.";
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL
CC DEATH.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76104; CAA53712.1; -
DR KSP; Q63450; 1A06.
DR MIM; 600831; -
DR InterPro; IPR000488; -
DR InterPro; IPR000719; -
DR InterPro; IPR002110; -
DR InterPro; IPR002290; -
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00069; kinase; 1.
DR PROSITE; PS50088; ANK_REPEAT; 6.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50017; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DM; 1.
DR PROSITE; PS50108; PROTEIN_KINASE_ST; 1.
KW Phosphatase; Serine/threonine-protein kinase; Calmodulin-binding;
KW phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
FT DOMAIN 13 266 PROTEIN KINASE.
FT DOMAIN 267 334 CALMODULIN-BINDING.
FT REPEAT 378 407 ANK 1.
FT REPEAT 411 440 ANK 2.
FT REPEAT 444 473 ANK 3.
FT REPEAT 478 507 ANK 4.
FT REPEAT 511 540 ANK 5.
FT REPEAT 544 573 ANK 6.
FT REPEAT 577 606 ANK 7.
FT REPEAT 610 639 ANK 8.
FT REPEAT 876 905 ANK 9.
FT REPEAT 1163 1197 ANK 10.
FT DOMAIN 1313 1397 DEATH.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).

```



```

RX MEDLINE-93104681; PubMed=1281700;
RA Zhang H., Scheirer D.C., Fowle W.H., Goodman H.M.;
RT "Expression of antisense or sense RNA of an ankyrin repeat-containing
RT gene blocks chloroplast differentiation in Arabidopsis.";
RL Plant Cell 4:1575-1588(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TEMPORAL AND SPATIAL REGULATION
CC OF CHLOROPLAST DEVELOPMENT FROM PROPLASTID.
CC -1- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION OCCURS IN TWO-WEEK-OLD
CC PLANTS AND DECLINES AS PLANTS DEVELOP FURTHER.
CC -1- SIMILARITY: CONTAINS 4 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M82883; AAA32812.1; -.
DR PIR; J01729; J01729.
DR HSSP; Q13625; 1YCS.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 4.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Cytoskeleton; Repeat; ANK repeat; Multigene family.
FT REPEAT 288 317 ANK 1.
FT REPEAT 321 350 ANK 2.
FT REPEAT 354 383 ANK 3.
FT REPEAT 387 416 ANK 4.
FT SEQUENCE 439 AA; 49150 MW; C371A90028B25BF3 CRC64;
-----
alignment_scores:
  Quality: 236.50      Length: 161
  Ratio: 2.319        Gaps: 2
Percent Similarity: 63.354 Percent Identity: 37.267

alignment_block:
US-09-509-775-1 x AKR_ARATH ..
Align seg 1/1 to: AKR_ARATH from: 1 to: 439

265 GAATTTTCTTG.....CAACTGGAGTCCAGTGAATGA 299
272 LysPheMetLeuAsnSerArgAsnProAspLeuAlaValAlaThrSerLy 288
300 TAAAGACGATGCAGGTGGTCTCTTCATATTGGCGCTTCGCTGGCC 349
288 slys.....TrpLeuProLeuHisThrLeuAlaAlaCysGlyC 301
350 GGGATGAGATTGTAAAGCCCTTCGGGAAAGGTGCTCAAGTGAATGCT 399
301 lupPheTyrLeuValAspSerLeuLeuLysHisAsnLeuAspLeuAsnAla 317
400 GTCAATCAAAATGGCTGCTACTCCCTTACATTATTCAGCTTCGNAACAG 449
318 ThrAspValGlyGlyLeuThrAlaLeuHisArgAlaAlaIleGlyLysLy 334
450 GCATGAGATCGCTGTCATGTGTACTGGAAGCGGGGCTAATCCAGATGCTA 499
334 sGlnAlaIleThrAsnTyrLeuLeuArgGluSerAlaAsnProPheVal 351
500 AGGACATTATGAGGTACAGACGACCGGCGACCAAGCGGTAAAC 549
351 euAspAspGluGlyAlaThrLeuMetHisTyrAlaValGlnThrAlaSer 367
550 TTGAAGATGATTCATCTCTCTGCTACTACAAGCATCCACAACATCCA 599
368 AlaProThrIleLysLeuLeuLeuTyrAsnAlaAspIleAsnAlaGl 384
600 AGACACTGAGGGTAAACACTCCTCTACACTTAGCCTGTGATGAGGAGAG 649

```

```

384 naspArgAspGlyTyrThrProLeuHisValAlaValGlnAlaArgAs 401
650 TGGAGAAGCAACAACTGGTGTCTCCCAAGGCAAGTATTACATGTAG 699
401 exAspIleValLysLeuLeuLeuLysGlyAlaAspIleGluValLys 417
700 AATAAAGAGAAAGACACCCCTGCAAGTGGCC 732
418 AsnLysAspGlyLeuThrProLeuGlyPheAla 428

seq_name: SwissProt_39:ANKH_CHRVI
seq_documentation_block:
ID ANKH_CHRVI STANDARD; PRT; 323 AA.
AC Q0527;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN HOMOLOG PRECURSOR.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochrochmatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D / ATCC 17899 / DSM 180;
RX MEDLINE-93300842; PubMed=8390993;
RA Dolata W.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA Cusanovich M.A.;
RT "Nucleotide sequence of the heme subunit of flavocytochrome c from
RT the purple phototrophic bacterium, Chromatium vinosum. A 2.6-kilobase
RT pair DNA fragment contains two multiheme cytochromes, a flavoprotein,
RT and a homolog of human ankyrin."
RL J. Biol. Chem. 268:14426-14431(1993).
CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13419; AAA23315.1; -.
DR HSSP; Q00420; IAWC.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 8.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Signal; ANK repeat; Repeat.
FT SIGNAL 1 28
FT CHAIN 29 323 ANKYRIN HOMOLOG.
FT REPEAT 31 60 ANK 1.
FT REPEAT 64 93 ANK 2.
FT REPEAT 97 126 ANK 3.
FT REPEAT 130 159 ANK 4.
FT REPEAT 163 192 ANK 5.
FT REPEAT 196 225 ANK 6.
FT REPEAT 229 258 ANK 7.
FT REPEAT 262 291 ANK 8.
FT SEQUENCE 323 AA; 33453 MW; 3976D34A9BD607DF CRC64;

```

```

alignment_scores:
  Quality: 225.00      Length: 219
  Ratio: 1.891        Gaps: 2
Percent Similarity: 54.338 Percent Identity: 29.680

alignment_block:
US-09-509-775-1 x ANKH_CHRVI ..

```



```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94479; AAB94646.1; -
DR HSSP: O00421; LAWC.
DR MGD: MGI:1195267; ILK.
DR InterPro: IPR000719; -
DR InterPro: IPR002110; -
DR Pfam: PF00023; ank; 3.
DR Pfam: PF00069; pkinase; 2.
DR PROSITE: PS50297; ANK_REPEAT; 1.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Repeat; ANK repeat.
FT REPEAT 33 62
FT REPEAT 66 95 ANK 1.
FT REPEAT 99 128 ANK 2.
FT REPEAT 180 212 PH-LIKE.
FT DOMAIN 193 446 PROTEIN_KINASE.
FT NP_BIND 199 207 ATP (POTENTIAL).
FT BINDING 220 220 ATP (POTENTIAL).
SQ SEQUENCE 452 AA; 51347 MW; E0BB30C52FB6865E CRC64;
```

alignment_scores:

```
Quality: 217.50 Length: 160
Ratio: 2.364 Gaps: 1
Percent Similarity: 57.500 Percent identity: 33.125
```

alignment_block:

US-09-509-775-1 x ILK_MOUSE

Align seg 1/1 to: ILK_MOUSE from: 1 to: 452

```
292 GTGAATGATAAGACGATGCGAGTGGTCTCTCTCTCATATTGCGGCTTC 341
      :::::  ::::  :::::  :::::  :::::  :::::  :::::  :::::
27  LeuAsnGlnGlyAspAspHisGlyPheSerProLeuHisTyrAlaCysAr 43
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
342 TGCTGGCGGGATGAGATTGTAAGAGCCCTCTCTGGGAAAGGTGCTCAAG 391
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
43  ggluglyArgSerAlaValValGluMetLeuIleMetArgGlyAlaArgI 60
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
392 TGAATGCTGTCATCAATGCGTCTACTCTCTTACATTATTCAGCTTCG 441
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
60  LeAsnValMetAsnArgGlyAspThrProLeuHisLeuAlaAlaSer 76
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
442 AAAAAGCAGCATGAGCTGCTCATGTTACTTGTGAAGGGGGGCTAATCC 491
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
77  HisGlyHisArgAspIle..... 82
492 AGATGCTAAGACCATATTAGGGTACAGCAATGCCGGGCGAGGCCA 541
82  ..... 82
542 AGGGTAAGTGAAGATGATTATATCTCTCTACTACAAAGCATCCACA 591
83  .....ValGlnLysLeuLeuGlnTyrFLysAlaAspThr 93
592 AACATCCAGACACTGAGGGTAACACTCTCTTACACTTAGCGCTGTGATCA 641
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
94  AsnAlaValAsnGluHisGlyAsnValProLeuHisTyrAlaCysPheTr 110
642 CGAGAGAGTGGNAGAGCAAACTGCTGTCCTCCCAAGGACGACTATT 691
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
```

```
110 pGlyClnAspGlnValAlaGluAspLeuValAlaAsnGlyAlaLeuValS 127
692 ACATTGAGAATAAAGAAAGAAAGACACCCCTGCAAGTGGCCAAAGGTGGC 741
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
127 erIleCysAsnLysTyrGlyGluMetProValAspLysAlaLysAlaPro 143
742 CTGGGTTTAAATACTCAAGAGAATGGTGGAA 771
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
144 LeuArgGluLeuLeuArgGluArgAlaGlu 153
```


OM of: US-09-509-775-1 to: SPTREMBL_16.* out_format : pfs

Date: Aug 13, 2001 8:36 AM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool/US09509775/runat_13082001_074528_13074/app_query.fasta_1.848  
-DB=SPTREMBL_16 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct  
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM-ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09509775_@CGN1_1_111 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY  
-WAIT -THREADS=1
```

Search information block:

Query: US-09-509-775-1

Query length: 780

Database: SPTREMBL_16.*

Database sequences: 425026

Database length: 132305027

Search time (sec): 108.000000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
sp_human:075832	+	1164.00	1881.43	6.4e-97	226	! 075832 homo sapiens (human). 26
sp_human:0922X3	+	1106.00	1787.22	1.1e-91	231	! 0922X3 rattus norvegicus (rat). 26
sp_human:0922X2	+	1105.00	1785.60	1.4e-91	231	! 0922X2 mus musculus (mouse). 26
sp_human:095533	+	582.50	943.92	2.0e-44	118	! 095533 homo sapiens (human). 26
sp_plant:092Q79	+	338.00	541.58	2.5e-22	247	! 092Q79 arabidopsis thaliana (ma
sp_human:092Q79	+	312.50	483.62	5.4e-20	194	! 092Q79 mus musculus (mouse). 26
sp_fungi:092P36	+	310.00	496.52	8.4e-20	237	! 092P36 neurospora crassa. relat
sp_human:013484	+	310.00	484.24	8.4e-20	208	! 013484 homo sapiens (human). 26
sp_human:012955	+	310.00	473.03	9.3e-20	437	! 012955 homo sapiens (human). 26
sp_human:010511	+	306.00	470.67	2.1e-19	262	! 010511 rattus norvegicus (rat)
sp_invertebrate:092767	+	305.00	480.46	2.5e-19	636	! 092767 caenorhabditis elegans
sp_invertebrate:088521	+	301.00	465.77	5.8e-19	1762	! 088521 rattus norvegicus (rat)
sp_invertebrate:09NCP8	+	297.00	462.65	1.3e-18	1159	! 09NCP8 drosophila melanogast
sp_invertebrate:09VSA2	+	297.00	456.65	1.4e-18	2443	! 09VSA2 drosophila melanogast
sp_human:09HAS4	+	295.50	459.52	1.7e-18	1265	! 09HAS4 homo sapiens (human). 26
sp_human:09H2K2	+	293.50	456.93	2.7e-18	1166	! 09H2K2 homo sapiens (human). 26
sp_mammal:09N180	+	290.50	452.28	5.1e-18	1136	! 09N180 bos taurus (bovine). 26
sp_human:013768	+	290.00	448.13	5.7e-18	1719	! 013768 homo sapiens (human). 26
sp_human:099407	+	290.00	447.51	5.8e-18	1856	! 099407 homo sapiens (human). 26
sp_invertebrate:017344	+	288.50	445.39	7.9e-18	1786	! 017344 caenorhabditis elegans
sp_invertebrate:017487	+	288.50	445.29	7.9e-18	1809	! 017487 caenorhabditis elegans
sp_invertebrate:017488	+	288.50	445.26	7.9e-18	1815	! 017488 caenorhabditis elegans
sp_invertebrate:017486	+	288.50	445.03	7.9e-18	1867	! 017486 caenorhabditis elegans
sp_invertebrate:017489	+	288.50	444.32	7.9e-18	2039	! 017489 caenorhabditis elegans
sp_invertebrate:017343	+	288.50	434.39	8.2e-18	6994	! 017343 caenorhabditis elegans
sp_invertebrate:061302	+	287.50	443.49	9.7e-18	1848	! 061302 mus musculus (mouse). 26
sp_invertebrate:097582	+	286.00	447.39	1.3e-17	843	! 097582 rattus norvegicus (rat)
sp_invertebrate:061304	+	285.50	444.45	1.4e-17	1098	! 061304 mus musculus (mouse). 26
sp_human:015084	+	283.50	441.50	2.2e-17	1059	! 015084 homo sapiens (human). 26
sp_human:095271	+	281.50	436.44	3.3e-17	1327	! 095271 homo sapiens (human). 26
sp_invertebrate:09XZ37	+	281.00	436.56	3.7e-17	1181	! 09XZ37 drosophila melanogast
sp_invertebrate:09VBP3	+	281.00	436.56	3.7e-17	1181	! 09VBP3 drosophila melanogast
sp_invertebrate:024241	+	276.50	427.08	9.5e-17	1549	! 024241 drosophila melanogast
sp_invertebrate:09V4B1	+	276.50	427.08	9.5e-17	1549	! 09V4B1 drosophila melanogast
sp_mammal:09TV77	+	270.50	422.53	3.2e-16	815	! 09TV77 sus scrofa (pig). 130 kb
sp_human:014974	+	270.50	420.64	3.3e-16	1030	! 014974 homo sapiens (human). 26
sp_invertebrate:062937	+	269.50	422.63	4.0e-16	658	! 062937 rattus norvegicus (rat)
sp_invertebrate:010728	+	269.50	419.46	4.0e-16	976	! 010728 rattus norvegicus (rat)
sp_invertebrate:090624	+	268.50	417.94	4.9e-16	963	! 090624 gallus gallus (chicken)
sp_invertebrate:090623	+	268.50	417.61	4.9e-16	1004	! 090623 gallus gallus (chicken)

sp_invertebrate:09XZC0 + 266.50 411.71 7.6e-16 1395 ! 09XZC0 latrodectus tredec
sp_human:09ULJ7 + 266.00 415.71 8.2e-16 768 ! 09ULJ7 homo sapiens (human).
sp_invertebrate:09VCA6 + 264.50 403.96 1.2e-15 2441 ! 09VCA6 drosophila melanog
sp_invertebrate:09VAU5 + 263.50 403.48 1.4e-15 2119 ! 09VAU5 drosophila melanog
sp_rhodent:089019 + 263.00 408.24 1.6e-15 1062 ! 089019 mus musculus (mouse)

seq_name: sp_human:075832

seq_documentation_block:

ID 075832 PRELIMINARY; PRT; 226 AA.

AC 075832;

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE 26S PROTEASOME SUBUNIT P28.

GN DJ889N15.2 OR HUMAN GANKYRIN.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98382579; PubMed=9714768;

RA Hori T., Kato S., Saeki M., DeMartino G.N., Slaughter C.A.,

RA Takeuchi J., Toh-E A., Tanaka K.;

RT "cDNA cloning and functional analysis of p28 (Nas6p) and p40.5

RT (Nas7p), two novel regulatory subunits of the 26S proteasome.";

RL Gene 216:113-122(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX Grafham D.;

RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RA Higashitsuji H., Fujita J.;

RT "Enhanced expression of a novel tumour marker in the human

RT hepatomas.";

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB009619; BAA33215.1; -.

DR EMBL; AL031117; CAA20117.1; -.

DR EMBL; D83197; BAA34594.1; -.

DR HSSP; P42773; 1IHB.

DR InterPro; IPR02110; -.

DR Pfam; PF00023; ank; 5.

DR PROSITE; PS50088; ANK_REPEAT; 5.

DR PROSITE; PS50297; ANK_REPEAT; 1.

DR SMART; SM00248; ANK; 1.

KW Proteasome.

SQ SEQUENCE 226 AA; 24428 MW; 57158E33146EC7C8 CRC64;

alignment_scores:

Quality: 1164.00 Length: 226

Ratio: 5.150 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-509-775-1 x 075832

Align seg 1/1 to: 075832 from: 1 to: 226

97 ATGAGGGGTGTGTCTTACCTTAATGGTCTGCAACCTGGCCCTACACGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG1 17

147 GAAGCTTGGAGAGCTTCAAGAGAGAGTATCTGGCCGATATAATCCCTGCTA 196
|||||
17 YlyLeuGluGluLeuLysLysSerLeuAlaAspLysSerLeuAla 34

197 CTGAAGACTGACACAGACAGACAGCAATGCAATTCACCTGGGATGCTCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

```
247 GCACATACAGAAATTTGTAATTTTGTGCAACTTGAGTGCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67
|||||
297 TGATAAAGCAGATGAGGTGGTCTCTCTTCAATATGCGGCTTCTGCTG 346
|||||
67 nAspLysAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84
|||||
347 GCCGGGATGAGATGTAAAGCCCTCTCGGAAAGGTGCTCAAGTGAAT 396
|||||
84 lYArgSpGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValasn 100
|||||
397 GGTGTCAATCAAAATGGTGTACTCTCTTACATTTATGAGTTCGAAAA 446
|||||
101 AlaValasnGlnAsnGlyCysThrProLeuHisIleAlaAlaSerLysas 117
|||||
447 CAGGCATGAGATCGGTGTCATGTTACTGGAAGGGGCTTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAlaSerAlaG 134
|||||
497 CTAAGGACCATTTATGAGCTACAGCAATGACCGGCGAGCCCAAGGTT 546
|||||
134 lAlLysAspHisTyGluAlaThrAlaMetHisArgAlaAlaLysGly 150
|||||
547 AACTTGAAGATGATTCATPATCTCTGTTACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetIleHisIleLeuLeuTyrrLysAlaSerThrAsnIl 167
|||||
597 CCAAGACACTGAGGTGATACCTCTCTACATTTAGCCTGTGATGAGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisIleAlaCysAspGluGlu 184
|||||
647 GAGTGAAGAAAGCAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTACAT 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyrrile 200
|||||
697 GAGATAAAGAAAGAAAGACCCCTGCAAGTGCCCAAGGTGCGCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuGl 217
|||||
747 TTTAATACCTCAGAGAAATGGTGAAGGT 774
|||||
217 yLeuIleLeuLysArgMetValGluGly 226
|||||
seq_name: sp_rodent:Q922X3
```

```
seq_documentation_block:
ID Q922X3 PRELIMINARY; PRT; 231 AA.
AC Q922X3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GANKYRIN HOMOLOGUE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022014; BAA36954.1;
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PSS0088; ANK_REPEAT; 5.
DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;
```

```
alignment_scores:
Quality: 1106.00 Length: 225
Ratio: 4.960 Gaps: 0
Percent Similarity: 99.111 Percent Identity: 94.667
alignment_block:
US-09-509-775-1 x Q922X3 ..
Align seg 1/1 to: Q922X3 from: 1 to: 231
97 ATGAGGGGTGTGTGTCTAAGTAACTGCTGCAACCTGGCTTACAGCGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrrAsnGl 17
|||||
147 GAAGCTCGAAGAGTTGAAGGAGAGTATCTGGCCGATAAATCCCTGGCTA 196
|||||
17 yLysLeuAspGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34
|||||
197 CTAGAAGTGAAGGAGGAGAGCAAGTGCATTTGCAACTGGGCTGCTCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50
|||||
247 GCACATCAGAAATTTGTAATTTTGTGCAACTTGAGTGCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67
|||||
297 TGATAAAGCAGATGAGGTGGTCTCTCTTCAATTTGCGGCTTCTGCTG 346
|||||
67 nGluLysAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84
|||||
347 GCCGGGATGAGATTTGTAAGCCCTCTCGGAAAGGTGCTCAAGTGAAT 396
|||||
84 lYArgSpGluIleValLysAlaLeuLeuLysGlyAlaGlnValasn 100
|||||
397 GCTGTCAATCAAAATGGTGTACTCTCTTACATTTATGCAAGTTCGAAAA 446
|||||
101 AlaValasnGlnAsnGlyCysThrAlaLeuHisTyrrAlaAlaSerLysas 117
|||||
447 CAGGCATGAGATCGGTGTCATGTTACTGGAAGGGGCTTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAlaSerProAspA 134
|||||
497 CTAAGGACCATTTATGAGCTACAGCAATGACCGGCGAGCCCAAGGTT 546
|||||
134 lAlLysAsnHisTyrrAspAlaThrAlaMetHisArgAlaAlaLysGly 150
|||||
547 AACTTGAAGATGATTCATPATCTCTGTTACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetValHisIleLeuLeuPheTyrrLysAlaSerThrAsnIl 167
|||||
597 CCAAGACACTGAGGTGATACCTCTCTACATTTAGCCTGTGATGAGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisIleAlaCysAspGluGlu 184
|||||
647 GAGTGAAGAAAGCAAACTGCTGTGTCCTCCCAAGGTGCGCTGGG 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyrrile 200
|||||
697 GAGATAAAGAAAGAAAGACCCCTGCAAGTGCCCAAGGTGCGCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuGl 217
|||||
747 TTTAATACCTCAGAGAAATGGTGAAGGT 771
|||||
217 yLeuIleLeuLysArgIleAlaGlu 225
|||||
seq_name: sp_rodent:Q922X2
seq_documentation_block:
ID Q922X2 PRELIMINARY; PRT; 231 AA.
AC Q922X2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
```

```

151 AsnLeuLysMetValHisIleLeuLeuPheTyrIleLysAlaSerThrAsnIle 167
597 CCAAGACACTGAGGGTAACACACCTCTACACTTACCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluGluA 184
647 GAGTGGGAAGCAAAACACTGCTGGTGTCTCCCAAGGAGCAAGTATTTACATT 696
|||||
184 rgValGluGluAlaLysPheLeuValThrGlnGlyAlaSerIleTyrIle 200
697 GAGAAATAAGCAAGAAAGACACCCCTGCAAGTGCACCAAGGTCGCTGGG 746
|||||
201 GlnAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG 217
747 TTTTAATACTCAAGAGAATGGTGGAA 771
|||||
217 yLeuIleLeuLysArgLeuAlaGlu 225

seq_name: sp_human:O95533

seq_documentation_block:
ID 095533 PRELIMINARY; PRT; 118 AA.
AC 095533:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)
DE DJ89N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
DE (PUTATIVE PARTIAL ISOFORM 2) (FRAGMENT).
GN DJ89N15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grahm D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL031177; CAA20118.1;
DR HSP; O13625; 1YCS.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 2.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
FT NON_TER 118
SQ SEQUENCE 118 AA; 12472 MW; F047E5F4C545D15A CRC64;

alignment_scores:
Quality: 582.50 Length: 151
Ratio: 4.936 Gaps: 1
Percent Similarity: 78.146 Percent Identity: 78.146

alignment_block:
US-09-509-775-1 x O95533 ..
Align seg 1/1 to: O95533 from: 1 to: 118

97 ATGAGGGGTGTGTCTAACCTAATGCTGTGCACCTGCCTACAGCGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG 17
147 GAAGCTGGAGAGTTCAGAGGAGATTTCTGGCGGATAAATCCCTGGCTA 196
|||||
17 yLysLeuGluGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34
197 CTAGAAGTCCAGGACAGACAGCAAGTGCATGTCACCTGGCATCTCAGCT 246
|||||
34 hrargThrAspGln..... 38
247 GGACATACAGAAATTTGTAATTTTGTGCACTTGGAGTCCAGTGAA 296
|||||
38 ..... 38

```

```

297 TCATAAGACGATGAGTGGTGTCTCTCTTCATATTCGGCTTCTGCTG 346
|||||
39 .....AlaGlyTTPSerProLeuHisIleAlaAlaSerAlaG 51
347 GCCGGGATGAGATGTTAAAGCCCTTCGGGAAGAGTGTCAAGTGAAT 396
|||||
51 lYArgAspGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValAsn 67
397 GCTGCAATCAAAATGGCTGTACTCCCTTACATTTAGCATTCAGCTTCGAAAAA 446
|||||
58 AlaValAsnGlnAsnGlyCysThrProLeuHisIleAlaAlaSerLysAs 84
447 CAGGATGAGATGCGTGTCTATCTTACTGGAAGCGGGCTTAATCCAGATG 496
|||||
84 nArghisGluIleAlaValMetLeuLeuGluGlyAlaAsnProAspA 101
497 CTAAGCATTATGAGCTACAGCAATGACCGGGGAGCCAGCCCAAGGTT 546
|||||
101 lAlYAspHisIleTyrGluAlaThrAlaMetHisArgAlaAlaAlaLysGly 117
547 AAC 549
|||
118 Asn 118
seq_name: sp_plant:Q9ZQ79
seq_documentation_block:
ID Q9ZQ79 PRELIMINARY; PRT; 247 AA.
AC Q9ZQ79;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PUTATIVE ANKYRIN.
GN T4W8.14.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T4W8 genomic sequence.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006284; AADI7433.1; -
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 247 AA; 26723 MW; 452884EBD70EFABE CRC64;

```

```

alignment_scores:
  Quality: 338.00      Length: 219
  Ratio: 2.315        Gaps: 6
Percent Similarity: 66.667 Percent Identity: 41.096

```

```
alignment_block:
```

```
US-09-509-775-1 x Q9ZQ79 ..
```

```
Align seg 1/1 to: Q9ZQ79 from: 1 to: 247
```

```

160 TTGAAGGAGAGATTTCTGCCCGATAAATCCCTGGCTACTAGACTGACCA 209
|||||
31 LeuSerGluGluGlnLeuSer...LysSerLeuAsnPheArg...AsnG 45

```

```

210 GGACAGCAGAACTCCATTGCATCGGGCATCGCTAGCTGGACATACAGAAA 259
|||||
45 uAspGlyArgSerLeuLeuHisValAlaAlaSerPheGlyHisSerGlnI 62
260 TTGTTGAATTTTGTG.....CAACTTGGAGTCCCAAGTGAATCAT 300
|||||
62 leValLysLeuLeuSerSerAspGluAlaLysThrValIleAsnSer 78
301 AAACAGCATGAGTGTGTCTCTCTTCTATTTGGCTTCTGCTGCTGCGCG 350
|||||
79 LysAspAspGluGlyTTPAlaProLeuHisSerAlaAlaSerIleGlyAs 95
351 GGATGAGATGCTAAAGCCCTTCTGGGAAAGTGTCAAGTGAATGCTG 400
: |||||
95 nAlaGluLeuValGluLeuLeuThrArgGlyValaAspValaAsnAla 112
401 TCAATCAAAATGGCTGTACTCCCTTACATTTAGCATTCAGCTTCGAAAAACAGG 450
|||||
112 ysAsnAsnGlyGlyArgThrAlaLeuHisIleAlaAlaSerLysGlyArg 128
451 CATGATGCGTGTCTATCTTACTGGAAGCGGGCTAATCCAGATGCTAA 500
|||||
129 LeuGluIleAlaGlnLeuLeuLeuThrHisGlyAlaLysIleAsnIleTh 145
501 GGACATTATGAGGCTACAGCAATGCACCGGGGAGCCAGCCCAAGGTAAC 550
|||
145 rAspLysValGlyCysThrProLeuHisArgAlaAlaSerValGlyLysL 162
551 TGAAGATGATTCAATCTCTCTGTACTACAAGCATCCACAACATCCAA 600
|||||
162 euGluValCysGluPheLeuIleGluGlyAlaGluIleAspAlaThr 178
601 GACACTGAGGCTAACACTCCCTCTA.....CACTTAGCTGTGATGAGGA 644
|||
179 AspLysMetGlyGlnThrAlaLeuMetHisSerValIleCysAspAspLys 195
645 GAGAGTG.....GAAGAAGCAAACTGCTGTGTCCTCCCAAG 679
: |||||
195 sGlnLeuLysValSerMetAspGlnValAlaPheLeuLeuIleArgHisG 212
680 GAGCAAGTATTTCATTCGATAAAGAAAGACACACCCCTGCAAGTG 729
|||||
212 lYAlaAspValAspValGluAspLysGlyGlyTyrThrValLeuGlyArg 228
730 GCCAAGTGCGCTGGGTTTAAATCTC.....AAGAGAATGTT 767
|||
229 AlaThrAsnGluPheArgProAlaLeuIleAspAlaAlaLysAlaMetLe 245
768 GGAAGGT 774
: |||||
245 uGluGly 247
seq_name: sp_rodent:Q61307
seq_documentation_block:
ID Q61307 PRELIMINARY; PRT; 1943 AA.
AC Q61307; Q61306; Q61308; Q61309; Q61310; O08866; O08867;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 3 (ANKYRIN G) (EPITHELIAL ANKYRIN) (ANKYRIN-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=95340633; PubMed=7615634;
RA Peters L.L., John K.M., Lu F.M., Elcher E.M., Higgins A., Vialamas M.,
RA Turtzo L.C., Olsuka A.J., Lux S.E.;
RT "Ank3 (epithelial ankyrin), a widely distributed new member of the

```

```

FT REPEAT 580 612 ANK MOTIF 18.
FT REPEAT 613 645 ANK MOTIF 19.
FT REPEAT 646 678 ANK MOTIF 20.
FT REPEAT 679 711 ANK MOTIF 21.
FT REPEAT 712 744 ANK MOTIF 22.
FT REPEAT 745 777 ANK MOTIF 23.
FT REPEAT 778 810 ANK MOTIF 24.
FT VARSPLC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLC 834 855 VRKASPEKISDGEVSDGEEG -> GDRCTWFKIPKVOEV
FT FT LVK (IN ISOFORM 5).
FT VARSPLC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND
FT FT ISOFORM 3).
FT VARSPLC 852 852 G -> GKCTWFKIPKVOEVLVKS (IN ISOFORM 4
FT FT AND ISOFORM 6).
FT VARSPLC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND
FT FT ISOFORM 7).
FT SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

alignment_scores:
  Quality: 312.50 Length: 203
  Ratio: 2.500 Gaps: 1
  Percent Similarity: 61.576 Percent Identity: 35.468

alignment_block:
US-09-509-775-1 x Q61307 ..

Align seg 1/1 to: Q61307 from: 1 to: 1943

220 ACTGCATTGCATGGCGCATGCTCAGCTGGACATACAGAAATGTTGAATT 269
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
517 ThrProLeuHisLeuAlaAlaArgGluGlyHisGluAspValAlaAlaPh 533

270 TTTGTTGCACTTGGAGTGGCCAGTGAATGATAAGACGATGAGTTGGT 319
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
533 eLeuLeuAspHisGlyAlaSerLeuSerIleThrThrLysLysGlyPheT 550

320 CTCCTCTTCATATTGGCGTCTGCTGCCGGGATGAGATTCTAAAGCC 369
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 hrProLeuHisValAlaAlaLysTyrGlyLysLeuGluValAlaSerLeu 566

370 CTTCTGGGAAAA..... 381
   ||||| |||

567 LeuLeuGlnLysSerAlaSerProAspAlaAlaGlyLysSerGlyLeuTh 583

381 ..... 381

583 rProLeuHisValAlaAlaAlaHisTyrAspAsnGlnLysValAlaLeuLeuL 600

382 ..... GGTCTCAAGTCAATGCTGTCANTCAAATGGCTGTACT 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 euLeuAspGlnGlyAlaSerProHisAlaAlaAlaLysAsnGlyTyrThr 616

421 CCCTTACATTATGCGCTTCGAAAAACAGGCATGAGATCGCTGTCATGTT 470
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633

471 ACTGGAAGCGGGCTAATCCAGATGCTTAAGCACCATATTAGGCTACAG 520
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
633 uLeuGluTyrGlyAlaAspAlaAsnAlaValThrArgGlnGlyIleAlaAs 650

521 CAATGTCACCGGGCAGCAGCAAGGTAACTTGAAGATGATTCATATCCTT 570
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
650 ervAlHisLeuAlaAlaGlnGluGlyHisValAspMetValSerLeuLeu 666

571 CTGTACTACAAAGCATCCACAAATCCAAAGACACTGAGGGGTAACACTCC 620
   ||| ..... ||||| ||||| ||||| ||||| ||||| |||||
667 LeuSerArgAsnAlaAsnValAlaAsnLeuSerAsnLysSerGlyLeuThrPr 683

621 TCTACACTTAGCCTGTGTAGCAGAGATGGAAGAAGCAAAACTGCTGG 670
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

380 AAGGTGCTCAAGTGAATGCTGTCAATCAAATGGCTACTACTCCCTTCACAT 4229
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
95 rgGlyAlaAspIleAsnGlnThrAsnHisGlnSerGlnThrAlaLeuHis 111
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
430 TATGCAGCTTCGAAACAACAGGCATGAGATCGCTGCTCATGTGTACTGGGAGG 479
    ::||| ||||| ||||| ||||| :::||||| ||||| ||||| |||||
112 PheIleAlaSerLysAsnAsnIleAspLeuAlaArgLysLeuLeuSerPr 128
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
480 CGGGGCTAATCCAGAT.....GCTAAGGACCATATTATCAGG 514
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||
128 oASpMetLysProLysProAlaSerValArgValLysAspLysArgGlyG 145
515 CTACAGCAATGACCGGGCAGCAGCCAAGGTAACTTGAAGATGATTCAT 564
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
145 lNtyrProLeuHisArgAlaAlaAlaIleGlySerValPrometIleasn 161
555 ATCCTTCGTACTACAAAGCATCCAAAACATCCAAACATCCAAACACTGAGGGTAA 614
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
162 LeuLeuLeuGlnHisLysSerProIleAsnAlaSerAspAsnAlaGlyTy 178
615 CACTCCTCTACACTTAGCTGCTGATGAGGAGAGAGTGGAAGAAGCAAAAC 664
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
178 rThrProLeuHisAlaValalagluGlyHisGlyHisAlalaVala 195
665 TGCTGCTGCTCCCAAGGACGCAAGTATT.....GCTAAGGACCATATTATCAGG 690
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
195 laLeuLeuLysAlaGlyAlaGluThrAspLysLysLysAspMetaspGlyTyr 211
691 .....TACATTGAGAA 701
212 LeuAlaLeuAspLeuAlaProaspLysAspValargArgPheIleGluTy 228
702 TAAAGAAGAAACACACCCCCTGCAAGTG 729
    :|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
228 sglulaGluLysgluglyleGluLeu 237

seq_name: sp_human:Q13484
seq_documentation_block:
ID Q13484 PRELIMINARY; PRT: 1088 AA.
AC Q13484;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN GL19.
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo-
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=96234670; PubMed=8666667;
RA Khasharian M., Morrow J.S., Ardito T., Stabach P.R., Mann I.L.,
RA Devarajan P.;
RA "Identification of a small cytoplasmic ankyrin (AnkGL19) in
RT and muscle that binds beta I sigma spectrin and associates
RT Golgi apparatus."
RL J. Cell Biol. 133:819-830(1996).
DR EMBL: U43965; AAB08437.1; -.
DR HSSP: P55273; 1B08
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 12.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 12.
DR PROSITE: PS50297; ANK_REP_REGION; 1.
DR SMART: SM00218; ZU5; 1.
SQ SEQUENCE 1088 AA; 119427 MW; 769C88D40A78DE86 CRC64;

alignment scores:

```



```

695 LysSerGlyLeuThrProLeuHisLeuAlaGlnGluAspArgValAs 711
654 AGAAGCAAAACTGCTGGTGTCCCAAGGACCAAGTATTACATTGAGAATA 703
711 nValAlaGluValLeuValAsnGlnGlyAlaHisValAspAlaGlnThrL 728
704 AGAAGAAAGAACACCCCTGCAAGTG 729
728 ysMetGlyTyrThrProLeuHisVal 736
seq_name: sp_rodent:070511

```

seq_name: sp_rodent:070511

seq_documentation_block:

```

ID 070511 PRELIMINARY; PRT; 2622 AA.
AC 070511;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1999 (TReMBLrel. 11, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE 270 KDA ANKYRIN G ISOFORM (ANKYRING) (FRAGMENT).
GN ANK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98417633; PubMed=9744885;
RA Zhang X., Bennett V.
RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
RT requires multiple ankyrin G-specific domains.*"
RL J. Cell Biol. 142:1571-1581(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Carpenter S.S., Zhang X.
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
RA Kordeli E., Ludsky M.A., Deprette C., Frappier T., Cartaud J.;
RL J. Cell Sci. 0:0-0(1998).
DR EMBL; AF102552; AAC78143.1; -
DR EMBL; AF065150; AAC18853.1; -
DR HSSP; P55273; IBD8.
DR InterPro; IPR000488; -
DR InterPro; IPR000906; -
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank_24; -
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
FT NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

```

alignment_scores:

```

Quality: 306.00 Length: 192
Ratio: 2.318 Gaps: 0
Percent Similarity: 68.750 Percent Identity: 36.458

```

alignment_block:

US-09-509-775-1 x 070511 ..

Align seg 1/1 to: 070511 from: 1 to: 2622

```

154 GAAGAGTTGAAGGACAGTATTCTGCGCCGCAATAATCCCTGGCTACTAGAAC 203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
553 GluAspValAlaAlaPheLeuLeuAspHisGlyAlaSerLeuSerIleTh 569

```

```

204 TGACCAGACAGACAGAACTGCATTGCACCTGGCATGCTCAGCTGGACATA 253
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
rThrLysLysGlyPheThrProLeuHisValAlaAlaLysTyrGlyLysL 586
254 CAGAAATGTTGAATTTTGTGCAACTTGAGCTGCCAGTCCCAATGATAAA 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
586 euGluValAlaSerLeuLeuGlnLysSerAlaSerProAspAlaAla 602
304 GACGATGCAGGTTGGTCTCCCTTCATATTGCGGCTTCTGCTGGCGCGGA 353
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
603 GlyLysSerGlyLeuThrProLeuHisValAlaAlaHisTyrAspAsnG 619
354 TGAGATTGTAAGCCCTTCTGGAAAAGGTGCTCAAGTGAATGCTGTCA 403
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
619 nLysValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 636
404 ATCAAAATGGCTGACTCCCTTACATTATGAGCTTCGAAAACAGGCAT 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
636 laLysAsnGlyTyrThrProLeuHisIleAlaAlaLysLysAsnGlnMet 652
454 GAGATCGCTGTCATGTTACTGGAAGGGGGCTAATCCAGATGCTAAGGA 503
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
653 AspIleAlaThrSerLeuLeuGluTyrGlyAlaAspAlaAsnProValTh 669
504 CCATTATGAGGCTACAGCAATGCACCGGGCAGCAGCCCAAGGGTAACTTGA 553
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
669 rArgGlnGlyIleAlaSerValHisLeuAlaAlaGlnGluGlyHisVala 686
554 AGATGATTATATCTCTGTACTACAAGCATCCACAACATCCAGAC 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
686 spMetValSerLeuLeuLeuSerArgAsnAlaAsnValAsnLeuSerAsn 702
604 ACTGAGGGTAACACTCCTCTACACTTAGCCCTGTGATGAGGAGAGTGGA 653
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
703 LysSerGlyLeuThrProLeuHisLeuAlaAlaGlnGluAspArgValAs 719
654 AGAAGCAAAACTGCTGGTGTCCCAAGGACCAAGTATTACATTGAGAATA 703
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
719 nValAlaGluValLeuValAsnGlnGlyAlaHisValAspAlaGlnThrL 736
704 AAGAGAAAGAACACCCCTGCAAGTG 729
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
736 ysMetGlyTyrThrProLeuHisVal 744
seq_name: sp_invertebrate:Q9T267
seq_documentation_block:
AC Q9T267 PRELIMINARY; PRT; 636 AA.
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE F40G9.1 PROTEIN.
GN F40G9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2; PubMed=7906398;
RX MEDLINE=94150718; Anderson K., Baynes C., Berks M., Coulson A.,
RA Wilson R., Ainscough R., Connel M., Copsey T., Cooper J., Fulton L.,
RA Bonfield J., Burton J., Durbin R., Favello A., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Hillier L., Jier M., Johnston L.,
RA Gardner A., Green P., Hawkins J., Kirsten J., Laister N., Latreille P.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Laister P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;

```


*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

```
RT RT elegans.
RL Nature 368:32-38(1994).
RN [2]
RC STRAIN-BRISTOL N2;
RA "The sequence of C. elegans cosmid F40G9.
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099919; AAC68798.1;
DR HSSP; Q00421; IAWC.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 636 AA; 73229 MW; AB55162AF5D5B1C0 CRC64;
```

alignment_scores:
Quality: 305.00 Length: 294
Ratio: 2.061 Gaps: 6
Percent Similarity: 50.340 Percent Identity: 29.252

alignment_block:
US-09-509-775-1 x Q9T267 ..

Align seg 1/1 to: Q9T267 from: 1 to: 636

```
109 GTGCTTAACCTAATGCTGTCGCAAC...CTGGCTACAGCGGGAAGCTGGA 155
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
335 ILELEUHIILELYSMETCYSLILELEUHIISTYRPHETHRGLYSASNVA 351
156 AGAGTTGAGGAGAGATGTTCTGCGCCGATAAATCCCTGGCTACTAGAACG 205
   ||| ||::: ::::: ::::: ::::: ::::: :::::
351 IGLUALALYSARGLEUETHRARGTYRPROLYSLEUVALGLITYRTHRA 368
206 ACAGGACAGCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
   ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
368 SPASPSERGLYARGSERTHRIEHIHPHEALALVALGLYGLYSERLEU 384
256 GAAATTGTTGAATTTTGTGCACTTGGAGTGCCAGTGAATGATAAGA 305
   ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
385 PROLEUETHRPHALALLE.....LEUASNASPPROGI 396
306 CGATGCA..... 312
396 uMetAlaHisLysThrAspValArgValGlyMetTrpAsnLeuThrg 413
312 ..... 312
413 luLysIleGluAsnLeuPhePheAsnArgLysIleCysAspPheArgLys 429
312 ..... 312
430 AsnTyrSerLysAsnThrLysAsnProHisPheSerGluPhePheGlu 446
312 ..... 312
446 uThrIleAspPheGluHisGluAsnSerGlnAsnAlaLeuLysLysPheP 463
312 ..... 312
463 heSerLeuLysIleAspIlePheAspLeuAsnCysLeuIleLeuProLeu 479
313 GGTGTGCTCTCTTCATATTGCGGCTTCTGCTGCGCGGATGAGATTGT 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
480 GlyTrpThrProLeuMetIleAlaSerSerAlaGlyArgValAspValVa 496
```

```
363 AAAAGCCCTTCTGGGA...AAAGGTGCTCAAGTGAATGCTGCTCAATCAA 409
   |::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
496 lArgTyrLeuLeuThrLeuProAspValAspValLysHisThrAsnSerA 513
410 ATGGCTGTACTCCCTTACATTATGCACTTCGAAACACAGGCATGAGATC 459
   || ||| ||||| ||||| ||||| ||||| ||||| |||||
513 snLysGlnThrSerLeuHisTyrAlaCysSerLysAsnHisValGluIle 529
460 GCTGTCATGTTACTGGAAGCGGGGCTAAT...CCAGATGCTAAGGACCA 506
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
530 VALLYSLEULEULEGUALAASPProAsnIleIleAsnLeuProAspLy 546
507 TTATGAGCTACAGCAATGCACCGGCAGCAGCAAGGTAAGTCACTGAAGA 556
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
546 spheGlyAlaThrAlaLeuHisArgAlaAlaSerArgGlyAsnAspValI 563
557 TGATTCATATCCTTCTGTACTAC...AAAGCATCCACAACATCCAGAC 603
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
563 lValArgAlaLeuValSerThrGlyLysCysSerLeuAspArgGlnAsp 579
604 ACTGAGGTTAACTCCTCTACACTTAGCTGTGATGAGGAGAGAGTGGA 653
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
580 GlyGluGlyAsnThrAlaLeuHisLeuAlaCysAspGluAsnArgGlyAs 596
654 AGAAGCAAAAACCTGCTGTCGCCAAGGAGCAAGTATTTACATTGAGAATA 703
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
596 pValAlaIleLeuLeuValAsnArgGlyAlaAspMetLysMetLeuAsnL 613
704 AAGAGAAAAGCACCCCTGCAAGTGCGCCCAA 735
   ||||| ::::: ||||| ||||| ||||| |||||
613 ysgLulysGlnThrProLeuGluMetLeuLys 623
```

seq_name: sp_rodent:088521

seq_documentation_block:

```
ID 088521 PRELIMINARY; PRT: 1762 AA.
AC 088521;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE 190 KDA ANKYRIN ISOFORM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kolli A.H., Devarajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, K-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1;
DR HSSP; P55273; 1BI8.
DR InterPro; IPR000169;
DR InterPro; IPR000488;
DR InterPro; IPR000906;
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
DR SMART; SM00005; DEATH; 1.
SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;
```

alignment_scores:
Quality: 301.00 Length: 192

```
Ratio: 2.263          Gaps: 0
Percent Similarity: 69.271  Percent Identity: 35.938
alignment_block:
US-09-509-775-1 x 088521  ..
Align seg 1/1 to: 088521 from: 1 to: 1762
154 GAAGAGTGAAGGAGAGTATTCTGCGCGGATAAATCCCTGCTACTAGAAC 203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
539 GluaspValAlaAlaPheLeuLeuAsHisGlyAlaPheLeuSerIleTh 555
204 TGACAGCAGCAGCAGAACTGCATTGCCTGGCGCATGCTCAGCTGGACATA 253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
555 rThrLysLysGlyPheThrProLeuHisValAlaAlaLysTyrgLysL 572
254 CAGAAATGTGAATTTTGTGTGCAACTTGGAGTCCGAGTGAATGATGATAA 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
572 euGluValAlaSerLeuLeuLeuGlnLysSerAlaSerProAspAlaAla 588
304 GACGATGCAGGTTGGTCTCCCTTCTCATATTGCGGCTTCTGCTGCGCGGA 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
589 GlyLysSerGlyLeuThrProLeuHisValAlaAlaHisTyraspAsnGI 605
354 TGAGATTGTAAAGCCCTTCTGGGAAAAGTGTCTCAAGTGAATGCTGTCA 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
605 nLysValAlaLeuLeuLeuLeuAspGlnGlyAlaSerProHisAlaAla 622
404 ATCAAAATGGCTGTACTCCCTTACATTATGCGCTTCGAAACACAGGCAT 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
622 laLysAsnGlyTyThrProLeuHisIleAlaAlaLysLysAsnGlnMet 638
454 GAGATCGCTGTACTTACTTGGGAGCGGGCTTAATCCAGATGCTTAAGGA 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
639 AspIleAlaThrSerLeuLeuGluTyrglyAlaAspAlaAlaValTh 655
504 CCATATGAGGCTACAGATGACCGCGGACGAGCCAGGAGGTAACTTGA 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
655 rArgGlnGlyIleAlaSerValHisLeuAlaAlaGlnGluGlyHisVala 672
554 AGATGATTTCATATCTTCTGTACTACAAAGCATCCACAAACATCCAGAC 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
672 spMetValSerLeuLeuLeuSerArgAsnAlaAsnValAsnLeuSerAsn 688
604 ACTGAGGCTAACACTCCCTCTACTTACCTGTGATGAGGAGAGAGTGGA 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
689 LysArgGlyLeuAsnProLeuHisLeuGlyGlyGlnGluaspArgValas 705
654 AGAAGCAAACTGTGTGTGTCGCCAGGAGCAAGTATTATTTACATTGAGAATA 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
705 nValAlaGluValLeuValAsnGlnGlyAlaHisValaspAlaGlnThrL 722
704 AAGAAGAAAGACACCCCTGCAAGTG 729
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
722 ysMetGlyTyThrProLeuHisVal 730
seq_name: sp_invertebrate:Q9NCP8
seq_documentation_block:
ID Q9NCP8 PRELIMINARY; PRT: 1159 AA.
AC Q9NCP8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE ANKYRIN 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The 11-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL; AF190635; AAF73309.1; -.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000906; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

alignment_scores:
Quality: 297.00      Length: 239
Ratio: 2.034         Gaps: 4
Percent Similarity: 61.088  Percent Identity: 29.707

alignment_block:
US-09-509-775-1 x Q9NCP8  ..
Align seg 1/1 to: Q9NCP8 from: 1 to: 1159
169 AGTATTCTGCGCGGATAAA.....TCCTGGCTACTAGACTGACCAGGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 SerLeuLeuLeuGluLysGlyGlyAsnIleGluAlaLysThr...Argas 270
213 CAGCAGAACTGCATTGCACCTGGCGCATCTCAGCTGGACATACAGAAATG 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 pGlyLeuThrProLeuHisCysAlaAlaArgSerGlyHisGluGlnVal 287
263 TTGAATTTTGTGCAACTTGGAGTGGCGTGAATGATGATAAGACGATGCA 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 alAspMetLeuLeuGluArgGlyAlaProIleSerAlaLysThrLysAsn 303
313 GGTGGTCTCTCT..... 324
|||:|||||
304 GlyLeuAlaProLeuHisMetAlaAlaGlnGlyGluHisValaspAlaAl 320
324 ..... 324
320 aaArgIleLeuLeuTyThrHisArgAlaProValaspGluValThrValasp 337
325 .....CTTCATATTGCGGCTTCTGCTGCGCGGATGAGATTGTA 363
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 yrLeuThrAlaLeuHisValAlaAlaHisCysGlyHisValArgValala 353
364 AAAGCCCTTCTGGGAAAAGGTGCTCAAGTGAATGCTGTCATCAAAATGG 413
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 LysLeuLeuLeuAspArgAsnAlaAlaAspAlaAsnAlaArgAlaLeuAsnGI 370
414 CTGTACTCCCTTACATTATGCGAGCTTCGAAAACAGGCATGAGATCGCTG 463
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 ypHeThrProLeuHisIleAlaCysLysLysAsnArgLeuLysValValG 387
464 TCATGTTACTGGAAGCGGGCTTAATCCAGATGCTAAGGACCATTTATGAG 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 luLeuLeuLeuArgHisGlyAlaSerIleSerAlaThrThrGluSerGly 403
514 GCTACGCAATGCGACCGGCGCAGCCAAAGGTAACCTTGAAGATGATTCA 563
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 LeuThrProLeuHisValAlaAlaPheMetGlyCysMetAsnIleValII 420
564 TATCCTTCTGTACTAAGACGATCCACAAACATCCACAGACACTGAGGTA 613
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 eTyThrLeuLeuGlnHisAspAlaSerProAspValProThrValArgGlyG 437
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:45:18 ; Search time 1165.93 Seconds
(without alignments)
6323.901 Million cell updates/sec.

Title: US-09-509-775-1
Perfect score: 780
Sequence: 1 tggtagagctctaacgctg.....gaatggtagaggttaaac 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est33:*
111: gb_est34:*
112: gb_est35:*
113: gb_est36:*
114: gb_est37:*
115: gb_est38:*
116: gb_est39:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuho, Y., Isogai, T.)
Unpublished (2000)
Contact: Takao Isoqai

**JOURNAL
COMMENT**

Description

SUMMARIES

Result No.	Score	Query			Description
		Match	Length	ID	
1	702	90.0	848	108	AU133088 AU133088
2	687	88.1	890	106	AL551457 AL551457
3	687	88.1	892	106	AL558577 AL558577
4	655.6	84.1	777	154	BG499627 602545002
5	646.4	82.9	773	173	BG107004 602991246
6	641	82.0	777	154	BG499442 602546667
7	620	79.5	867	140	BG783372 601471858
8	605.8	77.7	783	140	BE785737 601478508
9	604	77.4	634	155	BG568346 602587280
10	603.8	77.4	915	192	AK009068 Mus muscu
11	601	77.1	601	141	BE904296 601494565
12	597.8	76.6	646	136	BE540104 601061005
13	593.8	76.1	814	154	BG477395 602523604
14	590.6	75.7	829	153	BG426275 602492455
15	590.2	75.7	1301	192	AK018233 Mus muscu
16	584.8	75.0	854	146	BF211764 601812392
17	565.4	72.5	587	140	BE787558 601479443
18	564.8	72.4	856	141	BE893858 601436221
19	560.8	71.9	957	154	BG494082 602542388
20	555.8	71.3	786	140	BE787432 601479575
21	554.2	71.1	707	154	BG497031 602538951
22	552.2	69.7	860	174	BG175434 602337789
23	543.8	69.7	942	3	AA203428 zyx50e4.r
24	540.8	69.3	940	152	BG342503 602374061
25	528.8	67.8	647	146	BF212788 601814051
26	524.8	67.3	624	136	BE539557 601060269
27	513.4	65.8	595	173	BG087666 H31A1E11-
28	510	65.4	802	154	BG494451 602539438
29	505.4	64.8	910	144	BF122766 601761294
30	504.8	64.7	999	153	BG388982 602414741
31	503.8	64.6	1274	146	BF241523 601878408
32	503.4	64.5	895	168	BF698958 602127304
33	497.6	63.8	774	144	BF119890 601757668
34	493	63.2	600	31	AV593821 AV593821
35	481	61.7	886	174	BG118585 602348329
36	474	60.8	547	144	BF077522 227450 MA
37	469.8	60.2	550	173	BG085094 H3108E09-
38	464.6	59.6	745	146	BF217380 601885750
39	457.8	58.7	698	32	AV712330 AV712330
40	450	57.7	461	7	AA425037 zyx06a10.r
41	447.8	57.4	765	154	BG526676 602558119
42	436	55.9	818	154	BG350573 602560158
43	435.8	55.9	495	1	AA035825 mh17D04.r
44	435.4	55.8	652	119	AW682245 EST01037
45	434.6	55.7	1002	146	BF217727 601882552

ALIGNMENTS

RESULT	1
AU133088	
LOCUS	848 bp mRNA EST 24-OCT-2000
DEFINITION	AU133088 NT2RP4 Homo sapiens cDNA clone NT2RP4001258 5', mRNA sequence.
ACCESSION	AU133088
VERSION	AU133088.1 GI:10993627
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 848)
TITLE	Ota.T., Sugiyama.T., Ishii.S., Suzuki.Y., Saito.K., Yamamoto.J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and Isogai.T. HRI human cDNA project Ota.T., Sugiyama.T., Ishii.S. Suzuki.Y.,

```

|||||
Db 601 GGAGCAAGTATTACATTGAGATAAAGAGAAAGACACCCCTGCAAGTGGCCAAAGGT 660
|||||
QY 739 ggcctgggttaataactcaagagaagaatggtggaagggttaaca 780
|||||
Db 661 GGCTGGGTTTAATACTCAAGAGAATGGTGAAGGTTAAACA 702
|||||

RESULT 2
AL551457 AL551457 890 bp mRNA EST 16-FEB-2001
LOCUS AL551457 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI064YP22 5
DEFINITION prime, mRNA sequence.
ACCESSION AL551457
VERSION AL551457.1 GI:12889423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 255 a 177 c 221 g 236 t 1 others
ORIGIN

Query Match 88.1%; Score 687; DB 106; Length 890;
Best Local Similarity 100.0%; Pred. No. 2.4e-195;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 gaaatgaggggtgtgtcttaactaatggtctgaacctggcctcacagcggaagctg 153
|||||
Db 1 GAAATGAGGGGTGTGTCTTAACCTAATGCTGTGCAACCTGGCTACAGCGGAAGCTG 60
|||||
QY 154 gaagagtgaaggagattcttggccgataaataccctggctactagaactgaccaggac 213
|||||
Db 61 GAAGAGTTGAGGAGAGATATTCTGGCCGATTAATCCCTGGCTACTAGAACTGACCAGGAC 120
|||||
QY 214 agcagaactgcattgcactgggcatgctcagctggacatacagaaaattgtgaattttg 273
|||||
Db 121 AGCAGAACTGCATTGCCTGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTG 180
|||||
QY 274 ttgcaacttggagtgccagtgaaatgaataagacagatcaggttggtctctcttcatt 333
|||||
Db 181 TTGCAACTTGGAGTGCCAGTGAATGATAAAGACAGATCAGGTGGTCTCCTCTCATATT 240
|||||
QY 334 gcggcttctgtgcgcgggatgagattgttaaaagcccttctgggaaagtgctcaagt 393
|||||
Db 241 GCGGCTTCTGCTGCGCGGGATGAGATTGTAAGACCCCTTCGGGAAAGGTGCTCAAGTG 300
|||||

```

```

QY 394 aatgctgtcaatacaaatggctgtactcccttacattatgacgttcgaaaaaacagcat 453
|||||
Db 301 AATGCTGTCAATCAAAATGGCTGTACTCCCTTACATTATGACAGTTGAAAAACAGCAT 360
|||||
QY 454 gagatcgctgtcatgttacttggaaggcggggttaataccagatgctgaagaccattatgag 513
|||||
Db 361 GAGATCGCTGTCTATGTTACTTGAAGCGGGGCTTAATCCAGATGCTAAGGACCAATTATGAG 420
|||||
QY 514 gctacagcaatgcaccggcagcagcagcgaagggttaacttgaagatgattcatatccctctg 573
|||||
Db 421 GCTACAGCAATGACCCGGCAGCAGCCAGGGTAACTTGAAGATGATTATATATCCCTCTG 480
|||||
QY 574 tactacaagcatcccaacaacatccaagacactgagggtaaacactcctctacacttagcc 633
|||||
Db 481 TACTACAAAGCATCCACAACATCCAAGACACTGAGGTAACACTCCTCTACACTTAGCC 540
|||||
QY 634 tgtgatgagagagatggaagaagcaaaactcgtctggtgtcccaagaggaagaattattac 693
|||||
Db 541 TGTGATGAGGAGAGAGTGGAAAGAACAAACTGCTGTGTCTCCCAAGGAGCAAGTATTAC 600
|||||
QY 694 attgagaataaagaagaagacacacccctgcaagtggcgaagggtgacctgggtttaata 753
|||||
Db 601 ATTGAGAATAAAGAAGAAGAACACCCCTGCAAGTGCCCAAGGTGGCTGGGTTTAATA 660
|||||
QY 754 ctcaagagaatggtggaagggttaaca 780
|||||
Db 661 CTCAAGAGAATGGTGAAGGTTAAACA 687
|||||

RESULT 3
AL558577 AL558577 892 bp mRNA EST 16-FEB-2001
LOCUS AL558577 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ006YN18 5
DEFINITION prime, mRNA sequence.
ACCESSION AL558577
VERSION AL558577.1 GI:12903232
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 892)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 255 a 177 c 222 g 238 t
ORIGIN

Query Match 88.1%; Score 687; DB 106; Length 892;
Best Local Similarity 100.0%; Pred. No. 2.4e-195;

```


Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

94	gaatggaagggtgtgtgtctaaacctaatggtctgaacctggctctacagcgggaagctg	153
1	GAATGGAGGGGTGTGTCTAACTTAATGGCTGCAACCTGGCCCTACAGCGGAAGCTG	60
154	gaagagtggaaggagagattcttggccgataaattccctggtctactagaactgaccagga	213
61	GAAGAGTTGAAGGAGAGATTTCTGGCCGATAAATCCCTGGCTACTAGAACTGACCAGGAC	120
214	agcagaactgcatctgactcgggcatgctcagctggacatcacagaaattgtgaatttttg	273
121	AGCAGAATGCGAATGGCACTGGCGATGCTCAGCTGGCATACAGAAATTTGTGAATTTTGTG	180
274	tgcgaacttgagtgccagtgaaatgataaagacgactgcaggttggctcctctctcatatt	333
181	TTGCAACTTGGAGTGCACGATGATTAAGACGATGCAGGTGGTCTCTCTTCATATT	240
334	gcggcttctgtcggccgggattgagattgtaaaagcccttctggaaaaaggtctcaagtg	393
241	CGCGCTTCTGCTGCCGGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTG	300
394	aatcgtccaatcaaatggctgtactcccttacattatgcagcttcgaaaaacaggcat	453
301	AATGCTGTCAATCAAAATGGCTGTACTTCCCTTACATTATGCAGCTTCGAAAACAGGCAT	360
454	gagatcgcgtgtcatgttacttgaaggcggggttaataccagatgctaagaccattatgag	513
361	GAGATCGCTGTATGTTACTTGAAGGCGGGCTTAATCCAGATGCTAAGGACCATATGAG	420
514	gctcacagaatgacccgggcagcagcaagggttaacttgaagatgattcatcctctctg	573
421	GCTACAGCAATGCAACCGGGCAGCAGCAAGGGTAAGTTGAAGATGATTATATCTCTTG	480
574	tactacaaagcatccacaaactccaagacactgagggtaaacactcctctacacttagcc	633
481	TACTACAAAGCATCCACAACATCCAAGACACTGAGGGTAACACTCCTCTACACTTAGCC	540
634	tgtgatgagagagatggaagaagcaaaaactgctggtgtcccaaggagcaagtatttat	693
541	TGTGATGAGGAGAGTGGAAAGAGCAAAACTGCTGTGTCTCCCAAGGACGACGATTTTAC	600
694	attgagataaagagaaaaagacacccctcagtgcccaaggctggcctgggtttaaata	753
601	ATTGAGATAAAGAAAGAAAGACACCCCTGCAAGTGCCAAAGGTGGCTTGGTTTAATA	660
754	ctcagaagaattggtggaaggttaaca	780
661	CTCAAGACAATGTTGGGAAGGTTAAACA	687

RESULT 4

BG499627	767 bp	mrna	EST	27-MAR-2001		
LOCUS	602545002F1	NIH_MGC_60	Homo sapiens cDNA clone IMAGE:4667448	5',		
DEFINITION	mRNA sequence.					
ACCESSION	BG499627					
VERSION	BG499627.1	GI:13461144				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi					
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.					
TITLE	1 (bases 1 to 767)					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, ph.D.					
	Email: cgapbs-r@mail.nih.gov					
	Tissue Procurement: DCTD/DTP					
	cDNA Library Preparation: CLONETECH Laboratories, Inc.					
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
	DNA Sequencing by: Incyte Genomics, Inc.					

QY 680 gagcaagtattacattgagaataaagaagaagaaagacacccctgcaagtggcacaaggtg 739
 |||||
 Db 602 GAGCAAGTATTACATTGAGATTAGAGAGAGAGAAAGACACCCCTGCAAGTGGCCARAGGTG 661
 |||||
 QY 740 gccct-ggggttaatactcaagaagagtggtggaagggttaaca 780
 |||||
 Db 662 GCCTGGGGTTTAACTACTCAAGAGAATGGGTGGGAAGGTTCAAA 703
 |||||

RESULT 5
 BG107004 773 bp mRNA EST 30-JAN-2001
 LOCUS 602291246F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4386179 5',
 mRNA sequence.
 ACCESSION BG107004
 VERSION BG107004.1 GI:12600850
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 773)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LEM1068 row: k column: 12
 High quality sequence stop: 655.
 Location/Qualifiers
 1. .773
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4386179"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lymph; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."
 BASE COUNT 216 a 157 c 207 g 193 t
 ORIGIN

Query Match 82.9%; Score 646.4; DB 173; Length 773;
 Best Local Similarity 98.8%; Pred. No. 3.6e-183;
 Matches 672; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 102 ggggtgtgtctaacctaatgtgtcgaacctggcctacagcgggaagctggaagatt 161
 |||||
 Db 1 GGGGTGTGTCTAACTTAATGTGTCTGCAACCTGGCTACAGCGGGAAGCTGGAAGATT 60
 |||||
 QY 162 gaaggagatcttctggccgataaactccctggctactagaactgaccagacagcagaac 221
 |||||
 Db 61 GAGGAGAGATTCTGGCCGATAAATCCCTGGCTACTAGAACTGACGACGACGAGAAC 120
 |||||
 QY 222 tgcatcgactgggcatgctcagctggacatacagaaaattgttgattttgttgcaact 281
 |||||
 Db 121 TGCATTGCATGGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTGTGCAACT 180
 |||||
 QY 282 tggagtccaagtgaatgataaaagacagatgcaggttggtctcctctcatattcgccctc 341
 |||||
 Db 181 TGGAGTGCCAGTGAATGATAAAGACCATGATGAGTGGTCTCTCTTCATATTCCGGCTTC 240
 |||||

QY 342 tgctggccgggatgagattgttaaaagccctcttggaagagctgctcaagtgaatgctgt 401
 |||||
 Db 241 TGCTGGCCGGGATGAGATTGTGTAAGAGCCCTTCTGGGAAAGGCTGCTCAAGTGAATGCTGT 300
 |||||
 QY 402 caatcaaaatggctgtactcccttacattatcagcttcgaaaaacaggcatgagatcgc 461
 |||||
 Db 301 CAATCAAAATGGCTGTACTCCCTTACATTATGAGCTTGAAAAACAGGCATGAGATCGC 360
 |||||
 QY 462 tgctatgttactggaagcggggttaactcaagatcgaagaccattatgaggtctacagc 521
 |||||
 Db 361 TGTGATGTTACTTGGAGAGCGGGGCTAATCAGATGCTAAGGACCATATGAGGCTACAGC 420
 |||||
 QY 522 aatgcaccgggcagcagcgaaggttaacttgaagatgattcatatcctctctgtactacaa 581
 |||||
 Db 421 AATGCACCGGCGAGCAGCAAGGTAACTTGAAGATGATTATATCTCTTCTACTACAA 480
 |||||
 QY 582 agcatccacaaacatccaaagacactgagggtaacactctctacacttagcctgtgatga 641
 |||||
 Db 481 AGCATCCCAAAACATCCAAAGACACTGAGGGTAACCTCTCTACACTTACCTGTGATGA 540
 |||||
 QY 642 gtagagatggaagaagc-aaactgctgtgtgtcccaaggaagattattacattgaga 700
 |||||
 Db 541 GGAGAGAGTGGAGAGAACAAACACTGCTGTGTCCCAAGGAGCAAGTATTACATTGAGA 600
 |||||
 QY 701 ataaagaagaaagacacccctgcaagtggcacaaggtggcctgttggttttaatactaca 760
 |||||
 Db 601 AATAAGAGAAAAAGACACCCCTGCAAGTGCCCAAGGTGG-CTGGGTTTAAATACTCACGA 659
 |||||

QY 761 gaatgggtggaaggttaaca 780
 |||||
 Db 660 GAATTGTGGCCGGGTTAACA 679
 |||||

RESULT 6
 BG499442 777 bp mRNA EST 27-MAR-2001
 LOCUS 602546667F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4669048 5',
 mRNA sequence.
 ACCESSION BG499442
 VERSION BG499442.1 GI:13460959
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 777)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LEM1480 row: e column: 17
 High quality sequence stop: 639.
 Location/Qualifiers
 1. .777
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4669048"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: prostate; Vector: pDNK-LIB (Clontech);
 Site:1: SfiI (ggcgcctcgcc); Site:2: SfiI (ggccattatgccc
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor

FEATURES
 source

sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-DT(30)BN-3'
(where B = A, C, G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library.

BASE COUNT 218 a 162 c 213 g 184 t
ORIGIN

Query Match 82.2%; Score 641; DB 154; Length 777;
Best Local Similarity 97.3%; Pred. No. 1.5e-181;
Matches 684; Conservative 0; Mismatches 15; Indels 4; Gaps 3;

QY 80 gtctgggacacgaaatgaggggtgtgtctcaacctaatggtctgcaacctggcct 139
DB 3 GTTGTCTGGGACACGAAATGGAGGGGTGTGTCTAACCTTAATGGTCTGCAACCTGGCCT 62
QY 140 acagcgggaagctggaagagtgaagagagattcttgcgcgataaatccctggctacta 199
DB 63 ACAGCGGGAAGCTGGAAGAGTTGAAGGAGAGTATTCTGGCCGATAAATCCTCGCTACTA 122
QY 200 gaactgaccaggacagacagaactgcattgcaactggcattgctcagctggacatacagaaa 259
DB 123 GAACTGACCAGGACAGACAGAATGCATTGCACCTGGCATCTCAGCTGGACATACAGAAA 182
QY 260 ttgtgaattttgttgaacttggagtgccagtgatgaataaagacgatgcaaggttgt 319
DB 183 TTGTTGAATTTTTGTTGCAACTTGGAGTGGCCAGTGAATGATAAAGACGATGCGAGTTGGT 242
QY 320 ctctcttcataattgctgcttgcgcggagatgagattgttaaagccctctgggaa 379
DB 243 CTCCTCTCATATTGGCGCTTCTGCTGGCGGGATGAGATTGTAAGGCCCTTCCTGGGAA 302
QY 380 aagtgctcaagtgaatgtgttcaatacaaatggctgtactcctccctacattatgcagctt 439
DB 303 AAGTGCTCAAGTGAATGTGTCAATCAAAATGGCTGTACTCCTTACATTTATGCAGCTT 362
QY 440 cgaataacagcagatgagctgtctgtatgtactgaaagcgggtaataatccagatgcta 499
DB 363 CGAAACAGGCATGATGATGCTGTGTCTGTACTGGAAGCGGGGCTAATCCAGATGCTA 422
QY 500 aggaccattatgaggctacagcaatgcacccggcagcagcgaagggttaacttgaagatga 559
DB 423 AGGACCATTTAGAGCTACAGCAATGCACCGGCGACAGCCAGGCTTAATTCAGATGA 482
QY 560 ttcatatctctgtactacaaagcattccacaaacatccaaagacatgagggtaaacactc 619
DB 483 TTCTATATCTTCTGTACTACAAAGCATCCACAAACATCCAAAGACATGAGGGTAACACTC 542
QY 620 ctctacacttagctgtatgagagagagttgg-aagaagcaaaactgctgtgtctccaa 678
DB 543 CTCTACACTTAGCTGTGATGAGGAGAGTGTGAAGAAGCAAAACTGCTGGTGTCCCAA 602
QY 679 ggagcaagattttacattgagaataaagaagaagacacccctgcaagtggcgaaggt 738
DB 603 GGGAGCAAGTATTACATTGAGATATAAAGAGAA--GACACCTTGCAGTGGCCAGGGGT 660
QY 739 ggcctgggttttaatact-caagagaatggtggaaggttaaaaa 780
DB 661 GGCTGGGTTTAACTCCACGAGACTGGTGAAGGTAAACA 703

RESULT 7
BE783372 BE783372 867 bp mRNA EST 20-OCT-2000
LOCUS 601471858F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3874886 5',
DEFINITION mRNA sequence.
ACCESSION BE783372
VERSION BE783372.1 GI:10204570
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLN9633 row: c column: 15
High quality sequence stop: 653.

FEATURES
Location/Qualifiers
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3874886"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 253 a 158 c 246 g 210 t
ORIGIN
Query Match 79.5%; Score 620; DB 140; Length 867;
Best Local Similarity 98.6%; Pred. No. 3.2e-175;
Matches 657; Conservative 0; Mismatches 5; Indels 4; Gaps 3;
QY 112 tctaacctaatgctgcaacctgctcagcggaagctggaagctggaagagaggt 171
DB 1 TCTAACCTTAATGCTGTGCAACCTGGCCTACAGCGGAAGCTGGAAGAGTTGAAGAGAGT 60
QY 172 attctggcgcgataaatccctggctactgaaactgaccagacagcagaactgcattgcac 231
DB 61 ATTCTGGCCGATAAATCCTCGCTACTAGAACTGACCAGGACAGCAACTGCATTGCAC 120
QY 232 tgggcagtgcagctggacatacagaaaattgtggaattttgttgcgaacttggagtgcca 291
DB 121 TGGGCATGCTCAGCTGGACATACAGAAATGTTGAAATTTTGTGCAACTTGGAGTGCCA 180
QY 292 gtgaatgataaagacgatgcaggttggctcctcttcataattgcgcttctctggtccgg 351
DB 181 GTGAATGATAAGACCATGCGAGTTGGTCTCCTTTCATATTGCGGCTTCTGCTGCGCGG 240
QY 352 gatgagattgtaaaagcccttctgggaaaaggtgctcaagtgaatgctgtcaatcaaaat 411
DB 241 GATGAGATTGTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCTGTCAATCAAAAT 300
QY 412 ggcgtactccctacattatgcagcttcgaaaaaacagcagcatgagatcgtctgcatgta 471
DB 301 GGCTGTACTCCCTTACATTATGAGCTTCGAAAACAGGCATGAGATCGTGTGTCATGTTA 360
QY 472 ctggaagcgggggtaataatccagatcgaagaccattatgagctcacgaatgcaccgg 531
DB 361 CTGGAAGGCGGGCTAATCCAGATGCTAAGGACCATATGAGGCTACAGCAATGCAACCGG 420
QY 532 gcagcagcgaaggttaacttgaagatgattcattctctgttactacaagaatccaca 591
DB 421 GCAGCAGCCCAAGGTTAACTTGAAGATGATTTCATATCCTTCTGTACTACAAAGCATCCA 480
QY 592 aacatccaagacactgagggtaaacactctctacacttagcctgtgatgagagagtg 651
DB 481 AACATCCAAGACACTGAGGTAACACTCCTCTACACTTAGCTGTGATGAGGAGAGAGTG 540


```
/clone="2310001K15"
/sex="male"
/tissue.type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
24..719
/notes="putative"
/codon_start=1
/protein_id="BAB26053.1"
/db_xref="GI:12843632"
/translation="MEGCVSNIMCNLAYSGKDELKERILADKSLARTDDSTAL
HWACSGHTEIYEFLLQGVPNVNDKDGASPLHIAASAGDEIVKALVKGHAYNAV
NONGCTPLHYAASKRHEIYAVMLLEGANPDADHYDATAMHRAAKGNKMKVHILF
YKASNIQDEGTNPLHLACDEERVEEAKFLVTQASIVYENKEETPLQVAKGGLG
ILKRLAEGEASM"
polyA_signal 895..900
/notes="putative"
polyA_site 915
/notes="putative"
BASE COUNT 249 a 187 c 243 g 236 t
ORIGIN
Query Match 77.4%; Score 603.8; DB 192; Length 915;
Best Local Similarity 91.2%; Pred. No. 2.4e-170;
Matches 641; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 75 aagtagtctggagacagcaaatgaggggtgtgtgtctaaactaatggtctgaacct 134
DB 2 ACGGAGTCTTTGGGAGCGAAATGGAGGGGTGTGTCTAACATAATGATCTAACCT 61
QY 135 ggcctacagcgggaagctggaagagtggaagagtagcttctggccgataaatccctggc 194
DB 62 GGCCTACAGTGGGAAGCTGATGATGTGAAGAGCGCATTTTGGCTGATAAATCTCTGC 121
QY 195 tactagaactgaccagagacgacgaactgcattgcactggcgatgctcagctggacatac 254
DB 122 TACTAGAACTGATCAGGACAGCAGACAGCTTTGCACCTGGGCATGCTCAGCTGGCCATAC 181
QY 255 agaaattgtgaattttgtcaacttgagtcagtcagtcagtcagtcagtcagtcagtcagtc 314
DB 182 AGAATTGTTGAATCTTGTGCACTTGGAGTGGCAGTGCAGTGAATGATGAAGATGACGAGG 241
QY 315 ttgctctctctctaatattcggtctgtgctggtggtggtggtggtggtggtggtggtggt 374
DB 242 TTGCTCTCTCTTCAATATGCTGCTCGCTGCGCGGGATGAGATTGATAAGCCCTTCT 301
QY 375 gggaaaagggtcgaagtgaatgctgctcaatacaaatggtgttactccttactattatgc 434
DB 302 GGTGAAAGGTGCACATGATGATGCTGCAATCAAAACGGCTGCACCTCCACTCCATTATGC 361
QY 435 agcttcgaaaaacagcagatagatcgctgtcatgttactggaagcggtgcttaataccaga 494
DB 362 ACCTTCGAAGAATAGGCTAGATGCTGTATGTTACTAGAAAGTGGGGCTTAACCCAGA 421
QY 495 tgctaaggaccattatgagctacagcaaatgcaccggcgagcgcaagggtaacttgaa 554
DB 422 TGCAGAGACCATTTACGATGCTACAGCAATGCACCGGGCAGCAGCCAAAGGTTAACTTGA 481
QY 555 gatgattcatactcttctgactacaagaatcccaaaaatacaagaactgagggtaa 614
DB 482 GATGGTTACATCTCTTGTCTTCAAAAGGATCCACAAACATCCAAAGACACTGAGGGTAA 541
QY 615 cactctctacacttagctgtgagagagagtggtggaagaagcaaaactgctggtgc 674
DB 542 CACTCTCTACACTTACCTGTGATGAAGAGAGAGTGAAGAGAGCAAAATTTCTGTTGAC 601
QY 675 ccaaggagcaagtatttaccttgagaataagaagaagaagacacccctgcaagtggccaa 734
DB 602 TCAAGGAGCAAGTATTTACATTGAGATAAAGAGAAAGAAAGACACCCCTGCAAGTGGCAA 661
QY 735 aggtggcctgggttaatactcaagagaatggtggaaggttaa 777
|||||
```

```
Db 662 AGGGGGCCTGGGTTTAATACTCAAGAGACTAGCAGAAAGTGAA 704
RESULT 11
BE904296 601 bp mRNA EST 20-OCT-2000
LOCUS 601494565F2 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3896368 5',
DEFINITION mRNA sequence.
ACCESSION BE904296
VERSION BE904296.1 GI:10396403
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 601)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAN9689 row: b column: 17
High quality sequence stop: 599.
FEATURES
Location/Qualifiers
1..601
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3896368"
/clone_lib="NIH_MGC_70"
/tissue.type="epithelioid carcinoma"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 182 a 126 c 155 g 138 t
ORIGIN
Query Match 77.1%; Score 601; DB 141; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.4e-169;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 118 ctaatggtctgcaacctggcctacagcgggaagctggaagagtggaagagagattctg 177
DB 1 CTAATGGTCTGCAACCTGGCCTACAGCGGGAAGCTGGAAGAGTTGAAGAGAGATTCTG 60
QY 178 gccgataaatccctggctactagaactgaccaggaagcagaactgactgactgggca 237
DB 61 GCGGATAAATCCCTGGCTACTAGAACTGACCGAGCAGCAGCAACTGCATTGCCTGGCA 120
QY 238 tgcctagctggacacacagaaattgttgaattttgttgaacttggagtgccagtgaa 297
DB 121 TGTCTAGCTGGACATACAGAAATTTGTGAATTTTGTGCAACTTGGAGTGCAGTGAAT 180
QY 298 gataaagacgatgcaggtgtgtctctctctctctctctctctctctctctctctctctct 357
DB 181 GATAAAGACGATGCAGGTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 358 attgtaaaaaccccttctggaaaagggtgctcaagtgaatgctgcaatacaaatggctgt 417
DB 241 ATTGTAAAAACCCCTTCTGGGAAAAGGTGCTCAAGTGAATGCTCTCAATCAAAATGGCTGT 300
QY 418 actcccttacctatgatcgacttcgaaaaacagcagcatgagatcgctgtcatgttactgaa 477
DB 301 AC1CCCTTACATTATGCGAGCTTCAAAAACAGCATGAGATCGCTGTCTGTACTGTGAA 360
```

QY 478 ggcgggctaaatccagatgactaagaccattatgaggtacagcaatgcacccggcagca 537
 DB 361 GCGGGGCTAATCCAGATGCTAAGACCACTTATGAGGCTACAGCAATGCCCGGCAGCA 420
 QY 538 gcaagggttaactgaagatgattcattatccctctgtactacaaagcatcccaaacatc 597
 DB 421 GCGGGGCTAATCCAGATGCTAAGACCACTTATGAGGCTACAGCAATGCCCGGCAGCA 480
 QY 598 caagacactgaagggttaactcctctacacttagctgtgagagagagagtggaagaa 657
 DB 481 CAAGACACTGAGGGTAAACACTCCTTACACTTACACTTACACTTACACTTACACTT 540
 QY 658 gcaaaactgctgtgtcccaagagcaagatattacattgagaataaagaaagaaagaa 717
 DB 541 GCAAAACTGCTGTGTCTCCCAAGAGCAAGTATTTACATTGAGTAATTAAGAAAGACA 600
 QY 718 c 718
 DB 601 C 601

RESULT 12
 BE540104
 LOCUS 601061005F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3447625 5',
 DEFINITION mRNA sequence.
 ACCESSION BE540104
 VERSION BE540104.1 GI:9768749
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 646)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM8421 row: i column: 02
 High quality sequence stop: 570.
 Location/Qualifiers
 1..646
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3447625"
 /clone_lib="NIH_MGC_10"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5 kb. Library prepared by Life
 Technologies."

BASE COUNT 190 a 142 c 167 g 147 t
 ORIGIN

Query Match 76.6%; Score 597.8; DB 136; Length 646;
 Best Local Similarity 98.3%; Pred No. 1.4e-168;
 Matches 636; Conservative 0; Mismatches 7; Indels 4; Gaps 3;

QY 111 gtctaacctaatggtctgcaactgctgacagcgggagctggaagctggaagctggaagagag 170
 DB 1 GTCTAACCTAATGCTGCTGCAACCTGGCTACAGCGGGAAGCTGGAAGAGTTTCAAGAGAG 60

QY 171 tatttggccgataaattccctgctactagaaactgacaggagacagaaactgcatgca 230
 DB 61 TATTCCTGCGCGATAAATCCCTGGCTACTAGAACTGACAGGACAGCAGAACTGCATTGCA 120
 QY 231 ctgggcatgctcagctgacatcacagaaattgt-tgaatttttggcaacttgagtgcc 289
 DB 121 CTGGGCATGCTCAGCTGGACATACAGAAATTTCTGCAATTTCTGTTCGAACCTGGAGTGC 180
 QY 290 cagtaaatgataaagacagatgcaggttggtctcctctcattcattcggccttctgctggcc 349
 DB 181 CAGTGAATGATAAAGACGATGACAGGTGGTCTCTCTCTTCATATTGGGGCTTCTGCTGGCC 240
 QY 350 gggatgagattgaaaagccctcttgggaaaaggctgctcaagtgaatgctgcaatcaaa 409
 DB 241 GGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTGAATGCTGTCAATCAAA 300
 QY 410 atggctgtactcccttacattatgcagcttcgaaaacagggcatgagatcgctgtcatgt 469
 DB 301 ATGGCTGTACTCCCTTACATTATGCAATTCAGAGCTTCGAAAAACAGGCATGAGATCGCTCATGT 360
 QY 470 tactggaaggcggtgcttaatacagatgcttaagaccattatgaggtacagcaatgcacc 529
 DB 361 TACTGGAAGCGGGGCTAATCCAGATGCTAAGAGCACTTATGAGGCTACAGCAATGCACC 420
 QY 530 gggcagcagcccaagggttaactgaaagatgattcattcctctgtactacaaagcatcca 589
 DB 421 GGGCAGCAGCCCAAGGGTAACTTGAAGATGATTTCATATCCTTCTGTACTACAAGCATCCA 480
 QY 590 caaacatccagacactgagggtaaacactcctcctacacttagcctgtgtagggagagag 649
 DB 481 CAACATCCAGACACTGAGGGTAACTCCTCTACACTTAGCTGTGATGAGGAGAGAG 540
 QY 650 tgaagaagcaaaaactgctggtgtcccaaggagcaagtattacattgagaataaaa--ga 707
 DB 541 TGAAGAAGCAAAAACCTGTTGGTGTCCCAAGGAGCAAGTATCTACATTGAGAATACACGAA 600
 QY 708 agaaaagacacccctgcaagtggccaaaaggctggcctgggtttaataac 754
 DB 601 GCAAAAGACACCCCTGCAAGTGGCC-AAGGTGGCTGGGTTTAAATAC 646

RESULT 13
 BG477395 814 bp mRNA EST 21-MAR-2001
 LOCUS 602523604F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641840 5',
 DEFINITION mRNA sequence.
 ACCESSION BG477395
 VERSION BG477395.1 GI:13409674
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 814)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1409 row: h column: 01
 High quality sequence stop: 693.
 Location/Qualifiers
 1..814
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4641840"

FEATURES
 source

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:46:05 ; Search time 32.29 Seconds
(without alignments)
399,904 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AVSGKLEELKESILADKSLA.....TPLQVAKGGLILKRWEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDSL1/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1093	100.0	226	18 AAW15483	Human P28. Homo s
2	1093	100.0	226	20 AAY02430	Human gankyrin pro
3	1039	95.1	231	20 AAY02432	Rat gankyrin prote
4	1037	94.9	231	20 AAY02431	Mouse gankyrin pro
5	288.5	26.4	1166	22 AAB47022	Human SPANK. Homo
6	287	26.3	352	21 AAB11616	D. immitis ankyrin
7	287	26.3	1745	19 AAW70608	Full length ankyri
8	287	26.3	1745	19 AAW76776	D. immitis ankyrin
9	287	26.3	1745	21 AAB11589	D. immitis ankyrin
10	283.5	25.9	522	22 AAB66287	Human tankyrase2 c
11	283.5	25.9	1166	22 AAB66295	Human tankyrase2 T

12	283.5	25.9	1169	22 AAB66278	Human tankyrase2 r
13	283.5	25.9	1169	22 AAB66288	Human tankyrase2 c
14	283.5	25.9	1262	22 AAB66290	Human tankyrase2 c
15	283.5	25.9	1385	22 AAB66294	Human tankyrase2 T
16	281.5	25.8	673	21 AAY44403	Human truncated ta
17	281.5	25.8	949	21 AAY44404	Mouse SPANK. Mus
18	281.5	25.8	991	22 AAB47023	Human tankyrase I
19	281.5	25.8	1327	21 AAB27212	Human tankyrase I
20	281.5	25.8	1327	21 AAY44402	Human tankyrase I
21	281.5	25.8	1327	22 AAB66279	Human tankyrase I
22	281	25.7	1181	22 AAB66297	Drosophila tankyra
23	280.5	25.7	1166	22 AAY72589	Human tankyrase ho
24	277.5	25.4	1166	21 AAB27211	Human tankyrase II
25	271	24.8	302	19 AAW70609	Ankyrin protein PB
26	271	24.8	302	19 AAW76777	B. malayi ankyrin
27	271	24.8	302	21 AAB11590	B. malaya ankyrin
28	269.5	24.7	1030	19 AAW33572	Human myosin light
29	269.5	24.7	1030	19 AAW41378	Human protein p164
30	269	24.6	1074	20 AAY05734	Human Grb7 effecto
31	268.5	24.6	976	19 AAW53571	Rat p138 protein.
32	268.5	24.6	976	19 AAW41377	Rat protein p138.
33	265	24.2	756	22 AAB66286	Human tankyrase2 c
34	265	24.2	784	22 AAB66285	Human tankyrase2 c
35	265	24.2	907	22 AAB48574	Human breast cance
36	262	24.0	303	19 AAW70606	Ankyrin protein fr
37	262	24.0	303	19 AAW76774	D. immitis ankyrin
38	262	24.0	303	21 AAB11587	D. immitis ankyrin
39	260.5	23.8	978	21 AAB42288	Human ORFX ORF7052
40	252	23.1	763	21 AAY79154	Mouse protein kina
41	252	23.1	786	21 AAY69163	Amino acid sequenc
42	252	23.1	787	21 AAY76079	Murine protein kin
43	252	23.1	787	22 AAB6018	Skin cell protein,
44	250.5	22.9	378	22 AAB64963	Human secreted pro
45	250	22.9	982	19 AAW71632	Human myosin L-Cha

ALIGNMENTS

RESULT 1	
AAW15483	AAW15483 standard; Protein; 226 AA.
ID	AAW15483
XX	AAW15483;
AC	AAW15483;
XX	17-JUN-1997 (first entry)
DT	Human P28.
XX	Human; proteasome; P28; diagnosis; malignant tumour.
DE	Homo sapiens.
XX	JP09075085-A.
XX	25-MAR-1997.
PD	13-SEP-1995; 95JP-0235052.
XX	13-SEP-1995; 95JP-0235052.
PR	(SAGA) SAGAMI CHEM RES CENTRE.
XX	WPI; 1997-239267/22.
XX	N-PSDB; AAT66424-25.
XX	Human 26S proteasome constituting component protein - useful in the
PT	diagnosis of e.g. malignant tumour
XX	Claim 1; Page 6-7; 9pp; Japanese.
PS	This sequence represents the human proteasome component protein P28.
XX	The protein, P28, is useful for the diagnosis and treatment of
CC	

CC various diseases caused by proteasomes such as malignant tumour.

XX
SQ Sequence 226 AA; Best Local Similarity 100.0%; Score 1093; DB 18; Length 226; Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match
1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 aysgkleelkesiladkslatrtdqdsrtalhwacsaghteiveflqlgvpvndkddag 73
QY 61 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 120
DB 74 wslphiasagrdeivkallgkgaqvnavngctplhyaasknrheiaavmlleganpd 133
QY 121 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 akdhyeatamhraakgnlkmthillykastniqdtgntplhlacdeerveeakllvs 193
QY 181 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 213
DB 194 ggasiiyenkeektplqvakggglilkrmvveg 226

RESULT 2

AAV02430
ID AAY02430 standard; Protein; 226 AA.
XX
AC AAY02430;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
XX
OS Homo sapiens.
XX
PN WO9918201-Al.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR N-PSDB; AAX35852.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 1; Page 70-71; 11pp; Japanese.
XX
XX The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents human gankyrin.
XX
SQ Sequence 226 AA;

Query Match 100.0%; Score 1093; DB 20; Length 226;

Best Local Similarity 100.0%; Pred. No. 1.5e-111; Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 aysgkleelkesiladkslatrtdqdsrtalhwacsaghteiveflqlgvpvndkddag 73
QY 61 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 120
DB 74 wslphiasagrdeivkallgkgaqvnavngctplhyaasknrheiaavmlleganpd 133
QY 121 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 akdhyeatamhraakgnlkmthillykastniqdtgntplhlacdeerveeakllvs 193
QY 181 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 213
DB 194 ggasiiyenkeektplqvakggglilkrmvveg 226

RESULT 3

AAV02432
ID AAY02432 standard; Protein; 231 AA.
XX
AC AAY02432;
XX
DT 14-JUL-1999 (first entry)
XX
DE Rat gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
XX
OS Rattus sp.
XX
PN WO9918201-Al.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR N-PSDB; AAX35854.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 1; Page 76-78; 11pp; Japanese.
XX
XX The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents rat gankyrin.
XX
SQ Sequence 231 AA;

Query Match 95.1%; Score 1039; DB 20; Length 231;

Best Local Similarity 94.8%; Pred. No. 1.2e-105; Matches 201; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 aysgkleelkesiladkslatrtdqdsrtalhwacsaghteiveflqlgvpvndkddag 73

```
QY 61 WSPHIAASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 120
Db 74 WSPHIAASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
Db 134 aknydataamhraaakgnlkmvhillfykastniqdtgntplhlacdeerveeakllvt 193
QY 181 QGASIYIENKEEKTPLQVAKGGLGILKRWVE 212
Db 194 qgasiyienkeektplqvakggglgilkrlae 225

RESULT 4
ID AAY02431 standard; Protein: 231 AA.
AC AAY02431;
DT 14-JUL-1999 (first entry)
DE Mouse gankyrin protein.
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
OS Mus sp.
PN WO9918201-A1.
PD 15-APR-1999.
PF 02-OCT-1998; 98WO-JP04467.
PR 03-OCT-1997; 97JP-0286214.
PA (FUJI)/FUJITA.
PI Fujita J;
WPI; 1999-277266/23.
N-PSDB; AAX35853.
Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
PS Claim 6; Page 73-74; 11lpp: Japanese.
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents mouse gankyrin.
XX
SQ Sequence 231 AA;
Query Match 94.9%; Score 1037; DB 20; Length 231;
Best Local Similarity 93.9%; Pred. No. 2.1e-105;
Matches 199; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 AYSGLKEELKESILADKSLATRTDQSDRTALHWACSAGHTIVEFLQLGVPVNDKDDAG 60
Db 14 aysgkldeikeladkslatrtddqsdrtalhwacsaghtiveflqlgvpvndkddag 73
QY 61 WSPHIAASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 120
Db 74 WSPHIAASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
```

```
Db 134 aknydataamhraaakgnlkmvhillfykastniqdtgntplhlacdeerveeakflvt 193
QY 181 QGASIYIENKEEKTPLQVAKGGLGILKRWVE 212
Db 194 qgasiyienkeektplqvakggglgilkrlae 225

RESULT 5
ID AAB47022 standard; Protein: 1166 AA.
AC AAB47022;
DT 29-MAR-2001 (first entry)
DE Human SPANK.
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atherosclerosis; stroke;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 173..209 /note= "Encoded by nucleotides 768..878"
FT FT
FT Misc-difference 327..362 /note= "Encoded by nucleotides 1230..1337"
FT FT
FT Misc-difference 375..398 /note= "Encoded by nucleotides 1374..1445"
FT FT
FT Misc-difference 482..524 /note= "Encoded by nucleotides 1695..1823"
FT FT
FT Misc-difference 641..677 /note= "Encoded by nucleotides 2172..2282"
FT FT
W0200077225-A1.
21-DEC-2000.
09-JUN-2000; 2000WO-US15926.
11-JUN-1999; 99US-0138957.
(WHEED ) WHITEHEAD INST BIOMEDICAL RES.
(GEHO ) GEN HOSPITAL CORP.
Chi N, Lodish HF;
WPI; 2001-091404/10.
N-PSDB; AAC85294.
New insulin signalling protein SPANK, useful for reducing body mass,
glucose intolerance or insulin resistance and for preventing or
treating obesity-related and muscle-related diseases
Claim 3; Fig 3; 65pp; English.
This sequence represents human SPANK. The SPANK protein comprises
3 domains:
(a) a SAM (sterile alpha motif) domain;
(b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
domain; and
(c) an ANK domain composed of ankyrin repeats.
SPANK is a cytosolic protein which can poly(ADP-ribose)late itself.
SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
translocation of GLUT4 in the perinuclear region of adipocytes. It
is an effector in the insulin signalling pathway in eukaryotic cells.
```


(HESK-) HESKA CORP.

Blehm ES, Tang L;

WPI; 2000-375493/32.

N-PSDB; AAA58193, AAA58194, AAA58195, AAA58196.

New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding them, useful for treating and protecting animals from diseases caused by parasitic helminths, e.g. heartworm disease, elephantiasis or hydrocele

Example 1; Column 117-128; 120pp; English.

The invention relates to ankyrin proteins and nucleic acids from the parasitic helminths Dirofilaria immitis and Brugia malayi. It also relates to antibodies raised against such ankyrin proteins and to compounds that inhibit Dirofilaria or Brugia ankyrin function. Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour L3 cDNA library using PCR primers based on the sequence of the E1 ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids represent novel targets for anti-helminthic vaccines and drugs. Ankyrin nucleic acid molecules, proteins, vaccines and compositions are useful for protecting animals, particularly dogs, from diseases caused by parasitic helminths (e.g., heartworm disease, elephantiasis or hydrocele), as well as for treating the infection. The ankyrin nucleic acid molecules, proteins, vaccines and compositions of the invention are especially useful in treating and preventing infections caused by filarial nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria, strongylid, strongyloides, trichostrongyle, or trichurid nematodes and are also useful against cestodes and trematodes. The therapeutic compositions may be administered to mammals, including dogs, cats, humans, ferrets, horses, cattle, sheep, and other pets; economic food animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and compounds may also be used as diagnostic reagents to detect infection by parasitic helminths. Prior art anti-helminthic drugs require repeated administration, which often leads to the development of resistant helminth strains that no longer respond to treatment. Such drugs can also cause harmful side effects in the individual being treated, and a number of these drugs can only treat the symptoms of a parasitic disease, being unable to prevent infection by the parasitic helminth. Elucidation of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates the development of agents which inhibit ankyrin-mediated parasite developmental and migratory pathways. Sequence AAB11589 represents full-length D. immitis ankyrin, and sequences AAB11582-B11588, AAB11591-B11614 and AAB11616-B11629 represent D. immitis ankyrin fragments.

Sequence 1745 AA;

Query Match 26.3%; Score 287; DB 21; Length 1745;

Best Local Similarity 36.5%; Pred. No. 5.9e-22;

Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPNVDKDDA 59

Db 521 kegqeevaalmdhgtoktlitk---Kgtplhlaakynlpvaksllertgtpvdiiegkn 577

QY 60 GWSPLHTAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEIAVMLEGGANP 119

Db 578 qvtplhvaahyndkvalllengasahaaakngtptplhiaaknqmdiastillhykana 637

QY 120 DAKDHYEATAMHRAAAGNLMHILLYKASTNIQTEGNTPLHLACDERVEREAKLLV 179

Db 638 naeskagftplhlaaqeghremaallengakvgaqagrnngitpnhlcaqedrsvaeely 697

QY 180 SQGASIIYENKEETPLQVA 199

Db 698 kenaaidpktkagtyplhva 717

RESULT 10

AAB66287 AAB66287 standard; Protein; 522 AA.

XX AC AAB66287;

XX DT 05-APR-2001 (first entry)

XX DE Human tankyrase2 clone protein sequence SEQ ID NO: 99.

XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX OS Homo sapiens.

XX PN WO200100849-A1.

XX PD 04-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17827.

XX PR 29-JUN-1999; 99US-0141582.

XX PA (ICOS-) ICOS CORP.

XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX DR WPI; 2001-102896/11.

XX DR N-PSDB; AAF63925.

XX PT New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders

XX Example 1; Page 156-157; 242pp; English.

XX The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

Sequence 522 AA;

Query Match 25.9%; Score 283.5; DB 22; Length 522;

Best Local Similarity 32.1%; Pred. No. 2.4e-22;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPNVDKDDAGNS 62

Db 37 ngdverkrivtpkvnrsrdtagrkstplhfaagfgkrkdvveyllqnganvqardggli 96

QY 63 PLHTAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDAK 122

Db 97 plhnacsfghaeavnlllrhgadnrdwnytplhhaaakgkdkidvcivilqhgaetir 156

QY 123 D-----HYTEATAMHRAAAGN-LKMIHILLYKASTNIQDTEGNT 161

Db 157 ntgdrtaldladsakavltgkdkdelllesarsgnekmmalltpnvncshadgrkst 216

QY 162 PLHLACDEERVEAKLLVSOGASIIYENKEETPLQVA 199

Db 217 plhlaagynrvkvqlllqhgdvkhakdkgdlyplhna 254

RESULT 11

AAB66295 AAB66295 standard; Protein; 1166 AA.

XX AAB66295;
AC
XX
DT
XX
XX 05-APR-2001 (first entry)
DE Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.
XX
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW Inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US17827.
XX
XX 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
XX WPI: 2001-102896/11.
XX
XX N-PSDB; AAF63953.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX
XX Claim 3; Page 200-203; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.
XX
XX Sequence 1166 AA;
SQ

Query Match 25.9%; Score 283.5; DB 22; Length 1166;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SKLEELKESILADKSLATRTDODSRTALHWACSGAGTTEIVEFLQLGVPVNDKDDAGWS. 62
Db 34 ngdvervklvtpekvnsrdtagrktplhfaagfgrkdvveyllnganvqarddggll 93

QY 63 PLHIAASAGRDEITVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGGANPDAK 122
Db 94 plhnacsfghaeavnllrhgdpnrdnwnytplheaakgdkidvcivilqhgaeptir 153

QY 123 D-----HYEATAMHRAAKGN-LKMIHILLYKASTNIQDTEGNT 161
Db 154 ntldgrtaladpsakavltgeykdelliesarsgnekmmalltpnvnchsdgrkst 213

QY 162 PLHLCADERVEAKLLVSGASIIYENKEETPLQVA 199
Db 214 plhlaagynrvkivqlllqhgdadvhakdkgdlvplhna 251

RESULT 12
AAB66278
ID AAB66278 standard; Protein; 1169 AA.
XX
XX AAB66278;
AC
XX
XX 05-APR-2001 (first entry)
DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
DT

XX
DE Human tankyrase2 related protein sequence SEQ ID NO: 2.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW Inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US17827.
XX
XX 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
XX WPI: 2001-102896/11.
XX
XX N-PSDB; AAF63837.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX
XX Disclosure; Page 109-113; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.
XX
XX Sequence 1169 AA;
SQ

Query Match 25.9%; Score 283.5; DB 22; Length 1169;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SKLEELKESILADKSLATRTDODSRTALHWACSGAGTTEIVEFLQLGVPVNDKDDAGWS 62
Db 37 ngdvervklvtpekvnsrdtagrktplhfaagfgrkdvveyllnganvqarddggll 96

QY 63 PLHIAASAGRDEITVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGGANPDAK 122
Db 97 plhnacsfghaeavnllrhgdpnrdnwnytplheaakgdkidvcivilqhgaeptir 156

QY 123 D-----HYEATAMHRAAKGN-LKMIHILLYKASTNIQDTEGNT 161
Db 157 ntldgrtaladpsakavltgeykdelliesarsgnekmmalltpnvnchsdgrkst 216

QY 162 PLHLCADERVEAKLLVSGASIIYENKEETPLQVA 199
Db 217 plhlaagynrvkivqlllqhgdadvhakdkgdlvplhna 254

RESULT 13
AAB66288
ID AAB66288 standard; Protein; 1169 AA.
XX
XX AAB66288;
AC
XX
XX 05-APR-2001 (first entry)
DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW

KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX N-PSDB; AAF63926.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Example 1; Page 162-1665; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
XX Sequence 1169 AA;
SQ
Query Match 25.9%; Score 283.5; DB 22; Length 1169;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;
QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDAGWS 62
Db 37 ngdvervkrilvtpekvnsrdtagrkstplhfaagfgrkdvveyllqnganvqarddggli 96
QY 63 PLHTAASAGRDEIVKALLGKGAQVNAVNONGCTPLHYAASKNRHEIAVMLEGGANPDAK 122
Db 97 plhnacsfghaevvnlllrhgdapnrdnwnytpplheaakgkidvcivilqhgaepfir 156
QY 123 D-----HYEATAMHRAAAKGN-LKMIHILLYYKASTNIQDTGNT 161
Db 157 ntgdrtaldldpsakavltgcykddellesarsgnekmmalltpinvnchasdgrkst 216
QY 162 PLHLACDERVEEAKLLVSGQASIYIENKEEKTPLQVA 199
Db 217 plhlaagynrvkivqlilqhgdvhaakdkgdlvplhna 254
RESULT 14
AAB66290
ID AAB66290 standard; Protein; 1262 AA.
XX
AC AAB66290;
XX
XX 05-APR-2001 (first entry)
XX
XX Human tankyrase2 clone consensus protein SEQ ID NO: 107.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS

PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX N-PSDB; AAF63930.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Example 2; Page 173-176; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
XX Sequence 1262 AA;
SQ
Query Match 25.9%; Score 283.5; DB 22; Length 1262;
Best Local Similarity 32.1%; Pred. No. 8.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;
QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDAGWS 62
Db 130 ngdvervkrilvtpekvnsrdtagrkstplhfaagfgrkdvveyllqnganvqarddggli 189
QY 63 PLHTAASAGRDEIVKALLGKGAQVNAVNONGCTPLHYAASKNRHEIAVMLEGGANPDAK 122
Db 190 plhnacsfghaevvnlllrhgdapnrdnwnytpplheaakgkidvcivilqhgaepfir 249
QY 123 D-----HYEATAMHRAAAKGN-LKMIHILLYYKASTNIQDTGNT 161
Db 250 ntgdrtaldldpsakavltgcykddellesarsgnekmmalltpinvnchasdgrkst 309
QY 162 PLHLACDERVEEAKLLVSGQASIYIENKEEKTPLQVA 199
Db 310 plhlaagynrvkivqlilqhgdvhaakdkgdlvplhna 347
RESULT 15
AAB66294
ID AAB66294 standard; Protein; 1385 AA.
XX
AC AAB66294;
XX
XX 05-APR-2001 (first entry)
XX
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS
XX WO200100849-A1.
PN
XX PD 04-JAN-2001.
XX

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:45:21 ; Search time 19.4 Seconds
(without alignments)
226.069 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AVSGKLEELKESILADKSLA.....TPLOVAKGGLILKRMVEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	310	28.4	1088	4	US-09-082-059-2
2	294	26.9	1839	3	US-09-172-977-4
3	287	26.3	352	3	US-09-065-474-139
4	287	26.3	1745	2	US-09-031-485-33
5	287	26.3	1745	2	US-08-847-429A-33
6	287	26.3	1745	3	US-09-065-474-33
7	286	26.2	843	2	US-09-172-977-3
8	271	24.8	302	2	US-09-031-485-38
9	271	24.8	302	3	US-08-847-429A-38
10	271	24.8	302	3	US-09-065-474-38
11	262	24.0	303	2	US-09-031-485-23
12	262	24.0	303	2	US-08-847-429A-23
13	262	24.0	303	3	US-09-065-474-23
14	252	23.1	787	4	US-09-188-930-334
15	247	22.6	1423	4	US-08-810-712-10
16	241	22.0	348	2	US-09-031-485-28
17	241	22.0	348	2	US-08-847-429A-28
18	241	22.0	348	3	US-09-065-474-28
19	237	21.7	387	2	US-08-484-575A-7
20	237	21.7	387	3	US-08-477-459-7
21	237	21.7	387	3	US-08-479-869-7
22	237	21.7	387	4	US-08-486-414-7
23	237	21.7	387	5	PCT-US94-01826A-7
24	237	21.7	387	5	PCT-US94-02252A-7
25	227	20.8	752	1	US-08-281-193-2
26	227	20.8	752	1	US-08-422-106-2
27	227	20.8	752	2	US-08-735-716-2

28	227	20.8	752	2	US-08-555-568B-2	Sequence 2, Appli
29	227	20.8	752	5	PCT-US95-08069-2	Sequence 2, Appli
30	226.5	20.7	131	2	US-09-031-485-20	Sequence 20, Appl
31	226.5	20.7	191	2	US-08-847-429A-20	Sequence 20, Appl
32	226.5	20.7	191	3	US-09-065-474-20	Sequence 2, Appli
33	224.5	20.5	899	1	US-08-365-689-2	Sequence 2, Appli
34	224.5	20.5	899	1	US-08-145-138A-2	Sequence 2, Appli
35	224.5	20.5	933	1	US-07-747-781-2	Sequence 2, Appli
36	224.5	20.5	933	5	PCT-US92-06888-2	Sequence 2, Appli
37	218	19.9	394	2	US-08-555-568B-17	Sequence 17, Appl
38	218	19.9	687	2	US-08-555-568B-21	Sequence 21, Appl
39	218	19.9	688	2	US-08-555-568B-23	Sequence 23, Appl
40	211.5	19.4	452	3	US-09-035-706-2	Sequence 2, Appli
41	211.5	19.4	452	3	US-08-955-841-2	Sequence 2, Appli
42	210	19.2	657	1	US-08-264-534-34	Sequence 34, Appl
43	210	19.2	657	1	US-08-083-590A-13	Sequence 13, Appl
44	210	19.2	657	1	US-08-465-500-34	Sequence 34, Appl
45	210	19.2	657	2	US-08-346-128-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1

US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devorajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identifi
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082.059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 28.4%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 2.7e-26;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy	7	EELKESILADKSLATRTDSDSR	TALHWAC	SAGHTEI	VEFLQL	GLGVP	VNDKDD	AGWSP	LHI	66					
Db	166	EDVAFLDDHGASLITTKGFT	PLRVA	AKYK	LEVAN	LLQK	SASP	DACK	SGLT	PLHV 225					
Qy	67	AASAGRDEIKALIGKGAQ	NAV	QNGCT	PLHYA	AAK	NRH	EIA	VML	LEGGAN	PD	KDHYE 126			
Db	226	AHVDNQGKVALLLDQ	CAS	PHAAK	NGYTP	LHIAA	KKNQ	MDI	ATTT	LLEY	GADAN	AVTRQG 285			
Qy	127	ATAMHRAAAKGNLKM	II	LLYYK	ASTNI	QDT	EGT	PTPL	HAC	DEE	RV	VEAKLLV	SGGAS	IY 186	
Db	286	IASVHLAAQEGHVD	MSV	LL	LLGR	NAN	VLNS	KN	SGLT	PLHL	AAQ	EDRV	VNVAE	VLV	NOGAHD 345
Qy	187	IENKEEKT	PLQV	198											
Db	346	AQTKMGYTP	PLHV	357											

RESULT 2

US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.
APPLICANT: Yue, Henry
TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
FILE REFERENCE: PF-0615 US
CURRENT APPLICATION NUMBER: US/09/172,977
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 26.9%; Score 294; DB 2; Length 1839;
Best Local Similarity 31.9%; Pred. No. 4e-24;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;
QY 1 AVSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQ----- 48
DB 472 ARAGQVEVVR-CLLRNGALVDARAREEOTPLHIASRLGKTEIVQLLQHMHPDAATTNG 530
QY 49 -----LGVPNVDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVN 87
DB 531 YTPLHISAREGOVDVASVLEAGAAHSLATKGFPLHVAKYGSIDVAKLLQRRRAAD 590
QY 88 AVYNGCTPLHYAASKNRHEIATVLMLEGGANPDADKHDEATAMHRAAAKGNLKMTHILLY 147
DB 591 SAGKNGLTPLHYAAHYNDQKVALLLERKASPHATAKNGYTPHIAAKKNQMDIATLLN 650
QY 148 YKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIVYENKEEKTPLQVA 199
DB 651 YCAETNIVTKQGVTPHLASQEGHDMVTLLDKGANIHMSTKSLGTSLSHLA 702

RESULT 3
US-09-065-474-139
Sequence 139, Application US/09065474
Patent No. 6063599
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-09-065-474-139
Query Match 26.3%; Score 287; DB 3; Length 352;
Best Local Similarity 36.5%; Pred. No. 2e-24;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;
QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLQGVPNVDKDDA 59
DB 86 KEGQEEVAAILMDHGTDKTLTK---KGFTPLHIAAKYNLPVAKSLLERGTPTVDIEGKN 142
QY 60 GWSPLHIAASAGRDEIVKALLGKGAQVNQNGCTPLHYAASKNRHEIATVLMLEGGANP 119
DB 143 QVTPHLYAAHYNDKVALLLLENGASAAHAAKNGYTPHIAAKKNQMDIATLLHYKANA 202
QY 120 DAKDHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLV 179
DB 203 NAEKAGTTPHLAAGECHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRSVAAEELV 262
QY 180 SQGASIVYENKEEKTPLQVA 199
DB 263 KENAAIDPKTRAGYTPHVA 282

RESULT 4
US-09-031-485-33
Sequence 33, Application US/09031485
Patent No. 5824306
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,485
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER: US/08/847,429
FILING DATE: 24-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-031-485-33

Query Match 26.3%; Score 287; DB 2; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGAGQVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPHLHIAAKKNQMDIATLLHYKANA 637
QY 120 DAKDHYEATAMHRAAKGNLKMHIILLYKASTNTQDTGNTPLHLACDEERVEEAKLLV 179
DB 638 NAEKAGFTPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCQEDRVSVAEELV 697
QY 180 SOGASTIYIENKEEKTPLQVA 199
DB 698 KENAAIDPKTKAGYTPHLVA 717

RESULT 5

US-08-847-429A-33
Sequence 33, Application US/08847429A
Patent No. 5827692

GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heskia Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/847,429A
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-847-429A-33

Query Match 26.3%; Score 287; DB 2; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGAGQVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPHLHIAAKKNQMDIATLLHYKANA 637
QY 120 DAKDHYEATAMHRAAKGNLKMHIILLYKASTNTQDTGNTPLHLACDEERVEEAKLLV 179
DB 638 NAEKAGFTPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCQEDRVSVAEELV 697
QY 180 SOGASTIYIENKEEKTPLQVA 199
DB 698 KENAAIDPKTKAGYTPHLVA 717
RESULT 6
US-09-065-474-33
Sequence 33, Application US/09065474
Patent No. 6063599
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heskia Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-065-474-33

Query Match 26.3%; Score 287; DB 3; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGAGQVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPHLHIAAKKNQMDIATLLHYKANA 637

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/847,429A
;; FILING DATE: 24-APR-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Verser, Carol Talkington
;; REGISTRATION NUMBER: 37,459
;; REFERENCE/DOCKET NUMBER: HW-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-847-429A-38

Query Match 24.8%; Score 271; DB 2; Length 302;
Best Local Similarity 35.4%; Pred. No. 1e-22;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 11 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPLHIAASA 70
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIVYLLQOGANANVATVRGETPLHLAARA 116

Qy 71 GRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAKDHYEATAM 130
Db 117 NOTDIVRVLRNGAQVDAARELQTPHLIASRLGNTDIVILLQADASPNAATRDLYTL 176

Qy 131 HRAAAGNLMKMHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSQASIVYENK 190
Db 177 HIAAKEGEEVAAIILDHGSDKTLTKKGFTPLHLAAKYGNLPVAKLLERGLTLDIEGK 236

Qy 191 EEKTPLOVA 199
Db 237 NQVTPHVA 245

RESULT 10
US-09-065-474-38
; Sequence 38, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-065-474-38

Query Match 24.8%; Score 271; DB 3; Length 302;
Best Local Similarity 35.4%; Pred. No. 1e-22;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 11 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPLHIAASA 70
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIVYLLQOGANANVATVRGETPLHLAARA 116

Qy 71 GRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAKDHYEATAM 130
Db 117 NOTDIVRVLRNGAQVDAARELQTPHLIASRLGNTDIVILLQADASPNAATRDLYTL 176

Qy 131 HRAAAGNLMKMHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSQASIVYENK 190
Db 177 HIAAKEGEEVAAIILDHGSDKTLTKKGFTPLHLAAKYGNLPVAKLLERGLTLDIEGK 236

Qy 191 EEKTPLOVA 199
Db 237 NQVTPHVA 245

RESULT 11
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:

Query Match 24.0%: Score 262; DB 2; Length 303;

	Query	Match	Best Local Similarity	33.98;	Pred. No. 1.le-21;	Mismatches	97;	Indels	12;	Gaps									
QY	1	AYSCKLEELKESIIADKS	LA	RTQD	DS	-----	TALHWAC	SAGHTE	IV	EFLLQLG	VP	YND	55						
		:	:	:	:	:	:	:	:	:	:	:							
		:	:	:	:	:	:	:	:	:	:	:							
Db	16	ANCGHVRVAK	--	LLDRN	----	ADPNARALNG	F	PLHI	TACK	NNR	K	I	VELLLKYHA	IEA	69				
QY	56	KDDAGWSP	PLHIA	SAG	RDE	IV	EYK	ALLG	GAQV	NAV	NQNG	CTPL	HYA	AKSNR	HEIA	VM	LL	EG	115

```
Db 70 TTESGLSPLHVAEFGAINVIYLLQGANADVATVRGETPLHLAARANQTDIVRVLYRN 129
Qy 116 GANPAKDHYEATAMHRAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDEERVEA 175
Db 130 GAQVDAARELQTPHLIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAKEGQEEVA 189
Qy 176 KLVVSOGASIYIENKEETPLQV-AKGGGLGLILKRWVE 212
Db 190 AILMDHGTDKYLLTKGTTPPLHLAAKYGNLPVAKSILLE 227
```

```
RESULT 14
US-09-188-930-334
; Sequence 334, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 334
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-334
```

```
Query Match 23.1%; Score 252; DB 4; Length 787;
Best Local Similarity 31.4%; Pred. No. 6,5e-20;
Matches 64; Conservative 39; Mismatches 99; Indels 2; Gaps 2;

Qy 11 ESILADKSLATFTDSDSTALHWACSAGHTEIVEFLQLQGVVNDKDDAGWSPLHIAASA 70
Db 491 ELLARKTSVNAKDEQWTALHFAAQNGDEASTRLLEKNASVNEVDFEGRTPMHVACQH 550
Qy 71 GRDEIVKALLGGAOVNAVNGCTPLHYAASKNRHETAVMLL-EGGANPDAKDHYEATA 129
Db 551 GOENIVRTLRRGVVDVGLQGDWMLPLHYAAWQGHLPVKKLAKQPGVSVNAQTLDGRTP 610
Qy 130 MHRAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIYIEN 189
Db 611 LHAAQRGHYRVARLIDLCSDVNICSLOAQTPPLHVAETGHTSTARLLHLHGAGKEALT 670
Qy 190 KEETPLQV-AKGGGLGLILKRWVE 212
Db 671 SEGYPALHAAQNGHLATVAKLLIE 694
```

```
RESULT 15
US-08-810-712-10
; Sequence 10, Application US/08810712G
; Patent No. 6160106
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. LTD
; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
; TITLE OF INVENTION: Use of said Genes and Proteins
; FILE REFERENCE: sequencelist
; CURRENT APPLICATION NUMBER: US/08/810,712G
; CURRENT FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/11598
; EARLIER FILING DATE: 1994-10-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
```

```
; LENGTH: 1423
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-810-712-10

Query Match 22.6%; Score 247; DB 4; Length 1423;
Best Local Similarity 29.0%; Pred. No. 5,9e-19;
Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;

Qy 24 DDDSR TALHWACSAGHTEIVEFLQLQGVVNDKDDAGWSPLHIAASAGRDE----- 74
Db 401 DKGGSNAVYWAARHGHDVTLKPLSENKCPDLVDKSKGEMALHVAARYGHADVAQVTCAS 460
Qy 75 -----IVKALLCKGAOVNAVNGCTPLHYAASKNRHETA 109
Db 461 AQIPISRTKEETPLHCAAHGYYSVAKALCEAGCNVNIKNREGETPLLTASARGYHDIV 520
Qy 110 VMLLEGGANPDAKDHYEATAMHRAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDE 169
Db 521 ECLAERHAGADLNACDKDGHIALHLAVRRCOMIEVIKTLSSQGCFFVDYQDRHGNTPLHVACKD 580
Qy 170 ERVEEAKLLVSOGASIYIENKEETPLQVA 199
Db 581 GNMPIVWALCEANCNLDISNKYGRTPPLHLA 610

Search completed: August 13, 2001, 07:45:22
Job time: 59 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:46:40 ; Search time 22.54 seconds
(without alignments)
719.839 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AYSCKLEELKESILADKSLA.....TPLQVAKGGLGLIKRWVEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	338	30.9	247	D84448	probable ankryrin [
2	312.5	28.6	1765	T42714	ankryrin 3, splice
3	312.5	28.6	1940	T42715	ankryrin 3, splice
4	312.5	28.6	1943	T42713	ankryrin 3, splice
5	312.5	28.6	1961	T42716	ankryrin 3, splice
6	310	28.4	237	T50984	related to 26s pro
7	310	28.4	4377	T55575	ankryrin 3, long sp
8	304.5	27.9	636	T33631	hypothetical prote
9	294	26.9	3924	S37431	ankryrin 2, neurona
10	290.5	26.6	228	S57697	hypothetical prote
11	288.5	26.4	1786	A57282	ankryrin-related pr
12	288.5	26.4	1815	T15346	elegans ankryrin-re
13	288.5	26.4	1867	T15344	ankryrin-related un
14	288.5	26.4	2039	T15347	ankryrin-related un
15	287.5	26.3	1848	S37771	ankryrin, erythrocy
16	287.5	26.3	1862	T49502	ankryrin - mouse
17	286.5	26.2	1856	B35049	ankryrin 1, erythro
18	286.5	26.2	1880	A35049	ankryrin 1, erythro
19	286.5	26.2	1881	SJHUK	ankryrin 1, erythro
20	276	25.3	1549	T13940	ankryrin - fruit fl
21	273.5	25.0	234	T39032	hypothetical ankyr
22	269.5	24.7	815	JG0197	myosin-light-chain
23	268.5	24.6	658	S68418	protein phosphatas
24	267	24.4	1004	A55142	myosin-light-chain
25	260.5	23.8	1062	T30255	inversin - mouse
26	248	22.7	1062	T14151	inv protein - mous
27	247.5	22.6	1401	T11527	alpha-latrototoxin p
28	247	22.6	1423	I37275	death-associated p
29	245	22.4	397	T46445	hypothetical prote

30 240.5 22.0 1411 2 S30355
31 239.5 21.9 368 2 T18184
32 239 21.9 319 2 A57291
33 236.5 21.6 439 2 JQ1729
34 236.5 21.6 791 2 T42691
35 233 21.3 934 1 H71274
36 232 21.2 1398 2 T21884
37 231 21.1 557 2 T46507
38 229 21.0 1031 2 T43458
39 228 20.9 2352 2 T30201
40 226.5 20.7 1001 2 S30385
41 225.5 20.6 1188 2 T19552
42 225 20.6 323 2 B47169
43 224.5 20.5 900 2 A42024
44 224.5 20.5 933 2 S17233
45 222.5 20.4 851 2 T12503

alpha-latoinsecto
ankryrin repeat pro
cytokine inducible
ankryrin-repeat pro
hypothetical prote
probable ankryrin -
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
G9a protein - huma
ankryrin-like repea
transcription fact
transcription fact
hypothetical prote

ALIGNMENTS

RESULT 1

D84448

probable ankryrin [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: D84448

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: D84448

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <STO>

A:Cross-references: GB:AE002093; NID:g4335756; PIDN:AADI7433.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g03430

A:Map position: 2

Query Match 30.9%; Score 338; DB 2: Length 247;

Best Local Similarity 41.1%; Pred. No. 2.1e-22;

Matches 90; Conservative 32; Mismatches 81; Indels 16; Gaps 6;

QY 9 LKESILADKSLATRTDQDSRTALHWACSGAHTIEIVEFLL---QLGVPVNDKDDAGWSPLH 65

Db 31 LSEQLS-KSLNFR-NEDGRSLHLVAASFCHSQIVKLLSSSDEAKTVINSKDDGWAFLH 88

QY 66 IAASGRDEIVKALLGKAQVNAVNGCTPLHYAAKRNHETAVMLLEGGANPDADHY 125

Db 89 SAASIGNAEIVLLTRGDVNAKNGGRTALHYAAKGRLEIAQLLLTHGAKINITDKV 148

QY 126 EATAMHRAAKGNLKHITLLYYKASTNTQDTGNTPL--HLACDERV-----EEAKLL 178

Db 149 GCTPLHRAASVGLKEVCEFLIEEGAEIDATDKMGQTALMHSVICDDKQLKVSMDQVAFLL 208

QY 179 VSGQAGIYIENKEKTPLOVAKGGLGLIL----KRWVEG 213

Db 209 IRHGADVDDREGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 2

T42714

ankryrin 3, splice form 2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42714

R:Peters, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Vialamas, M.; Turtzo

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42714

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match 28.6%; Score 312.5; DB 2; Length 1765;

Best Local Similarity 35.5%; Pred. No. 4.2e-19;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHAAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAAYKGLVAVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGNTPLHLACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLVSGAGSIYIENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 3

T42715

ankryrin 3, splice form 3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42715

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42715

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1940 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 834/1

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match

Best Local Similarity 28.6%; Score 312.5; DB 2; Length 1940;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHAAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAAYKGLVAVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGNTPLHLACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLVSGAGSIYIENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 4

T42713

ankryrin 3, splice form 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42713

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 855/1

C:Function:

A:Description: supposed to play an important role in the polarized distribution of ma

A:Note: major kidney ankryrin

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match

Best Local Similarity 28.6%; Score 312.5; DB 2; Length 1943;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHAAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAAYKGLVAVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGNTPLHLACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLVSGAGSIYIENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 5

T42716

ankryrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42716

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42716

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1961 <PET>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 28.6%; Score 312.5; DB 2; Length 1961;
Best Local Similarity 35.5%; Pred. No. 4.8e-19;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

Qy 29 TALHWACSGHTEIVFEFLQLGVPVNDKDGWSPHHTAASAGRDEIVKALGK----- 82
Db 517 TPLHLAAREGHEDVAFLDHRGASISITTKGFTPLHVAAYKGVKLEVASLLQLKASPSDA 576
Qy 83 -----GAQVNAVNGCTPLHYAASKNRHEIAVMLEGG 115
Db 577 AGKSGLTPLHVAHYDNQKVALLLDQGASPHAAKNGYTPHIAAKKNQMDIATSLLEY 636
Qy 116 GANPAKDHYEYATAMHRAAKGNLMIHLLYYKASTNIQDTGEGTTPHLACDEERVEEA 175
Db 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696
Qy 176 KLVVSGASIVYENKEETPLQV 198
Db 697 EVLVNQGAVHDQTKMGYTPHLHV 719

RESULT 6
T50984
related to 26s proteasome subunit p28 [imported] - Neurospora crassa
N:Alternate names: protein B7F18.30
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
A:Accession: T50984
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: 225286
A:Status: preliminary
A:Accession: T50984
A:Molecule type: DNA
A:Residues: 1-237 <SCH>
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.30
A:Map position: 6
A:Introns: 17/1; 25/3; 68/3

Query Match 28.4%; Score 310; DB 2; Length 237;
Best Local Similarity 34.5%; Pred. No. 5.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

Qy 1 AVSGKLEELKESILADKSLATRTDQDSSTALHWACSGHTEIVFEFL-OLGVPVNDKDDA 59
Db 12 ARDGRASIVESLLNANPKLAQKDDGRLPIHWACSYNRKVEVLLVKNQKDFDVEDDM 71
Qy 60 GWSPLHIAASA-GRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 118
Db 72 GWTTPMISASVKSDAIIDLLSRGADINQTNHQSQTALHTASKNNDIDLAKKLSPOMK 131
Qy 119 PD-----AKDHYEATAMHRAAKGNLMIHLLYYKASTNIQDTGEGTTPHLACDEERVE 173
Db 132 KPASVRYKRGQYPLHRAAAGVSPINLLQLHQSINASDNAGYTPPLHVAEGHGH 191
Qy 174 EAKLVVSGASI-----YENKEETPLQV 198
Db 192 AAVALLKAGATDKKMDGYLALDLAPDKDVRRTFEKEAREGIEL 237

RESULT 7
A55575
ankyrin 3, long splice form - human
N:Alternate names: ankyrin G
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C:Accession: A55575
R:Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the
A:Reference number: A55575; MUID:95138209
A:Accession: A55575
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025
C:Genetics:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465
A:Map position: 10q21-10q21
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
C:Keywords: alternative splicing; peripheral membrane protein

F:73-105/Domain: ankyrin repeat homology <AN01>
F:106-138/Domain: ankyrin repeat homology <AN02>
F:139-171/Domain: ankyrin repeat homology <AN03>
F:172-200/Domain: ankyrin repeat homology <AN04>
F:201-233/Domain: ankyrin repeat homology <AN05>
F:234-286/Domain: ankyrin repeat homology <AN06>
F:267-299/Domain: ankyrin repeat homology <AN07>
F:300-332/Domain: ankyrin repeat homology <AN08>
F:333-365/Domain: ankyrin repeat homology <AN09>
F:366-398/Domain: ankyrin repeat homology <AN10>
F:399-431/Domain: ankyrin repeat homology <AN11>
F:432-464/Domain: ankyrin repeat homology <AN12>
F:465-497/Domain: ankyrin repeat homology <AN13>
F:498-530/Domain: ankyrin repeat homology <AN14>
F:531-563/Domain: ankyrin repeat homology <AN15>
F:564-596/Domain: ankyrin repeat homology <AN16>
F:597-629/Domain: ankyrin repeat homology <AN17>
F:630-662/Domain: ankyrin repeat homology <AN18>
F:663-695/Domain: ankyrin repeat homology <AN19>
F:696-728/Domain: ankyrin repeat homology <AN20>
F:729-761/Domain: ankyrin repeat homology <AN21>
F:762-794/Domain: ankyrin repeat homology <AN22>
F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match 28.4%; Score 310; DB 2; Length 4377;
Best Local Similarity 37.0%; Pred. No. 2.2e-18;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 7 EELKESILADKSLATRTDQDSSTALHWACSGHTEIVFEFLQLGVPVNDKDDAGNSPLHI 66
Db 545 EDVAFLDHRGASISITTKGFTPLHVAAYKGVKLEVASLLQLKASPSDAAGKSGLTPLHV 604
Qy 67 AASGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
Db 605 AAHYDNQKVALLLDQGASPHAAKNGYTPHIAAKKNQMDIATTLLEYGADANAVTRQG 664
Qy 127 ATAMHRAAKGNLMIHLLYYKASTNIQDTGEGTTPHLACDEERVEEAKLVVSGASIV 186
Db 665 TASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 724
Qy 187 IENKEETPLQV 198
Db 725 AQTGMGYTPHLHV 736

RESULT 8
T33631
hypothetical protein F40G9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

```

Query Match      26.9%; Score 294; DB 2; Length 3924;
Best Local Similarity 31.9%; Pred. No. 4.8e-17;
Matches      74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

QY      1  AYSKGLLEKESILADKSLATFTDODSRTALHWACSAGHTEIVEFLQ----- 48
DB      472  ARAGQVEVVR-CLLRNGALVDARREEQTPLHIASRLGTEIVQLLQRMHPDAATTNG 530
QY      49  -----LGVPNVDKDDAGVSGPLHIAASAGSAGDEIVKALLGKGAQVN 87
DB      531  YTPPLHISAREGOVDVASVLEGAHAHSLATKGFTPLHVAAKYGSGLDVAKULLQRRRAAD 590
QY      88  AVNQNGCTPLHYAASKNRHEIAVMLEGGANPDADKHYEATAMHRAAAKGNKMTIHILLY 147
DB      591  SAGKNGLTPLHVAAHVDNQKVALLELLEKASPHATAKNGYTPLHIAAKKNQMOIASTLLN 650
QY      148  YKASNTIQDTGENTPLHLACDEERVEEAKLLVYSQGASIIYENKEEKTPLQVA 199
DB      651  YGAETNIVTKQGTPLHLASQEGHDTMVTLLLDKGANIHMTSKSGLTSLHLA 702

RESULT      10
S57697

```


hypothetical protein YGR232w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G8564
C:Species: Saccharomyces cerevisiae
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Mar-2001
C:Accession: S57697; S64556; S63913
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 kb CRM1-YLM9-PET54-SM11-PHO81-YHB4-PFK1 region
A:Reference number: S57680
A:Accession: S57697
A:Molecule type: DNA
A:Residues: 1-228 <VAM>
A:Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61182.1; PID:g886926
R:van der Aart, Q.J.M.; Steensma, H.Y.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64541
A:Accession: S64556
A:Molecule type: DNA
A:Residues: 1-228 <VAM>
A:Cross-references: EMBL:Z73017; NID:g1323418; PIDN:CAA97260.1; PID:e243668; PID:g132341
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
Yeast 12, 385-390, 1996
A:Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-PFK1 region
A:Reference number: S63896; MUID:96267763
A:Accession: S63913
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-228 <VAF>
A:Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61182.1; PID:g886926
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 7R
A:Note: YGR232w
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:71-103/Domain: ankyrin repeat homology <ANI>

Query Match 26.68; Score 290.5; DB 2; Length 228;
Best Local Similarity 38.6; Pred. No. 2.9e-18;
Matches 76; Conservative 30; Mismatches 80; Indels 11; Gaps 5;
Qy 5 KLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLQGVVNVN---DKDDAGW 61
Db 17 KVQELLHS---KPSLLQKQDQGRPLHWSVFOAHEITSFLLSKMENVNLDYPPDDSGW 73
Qy 62 SPLHTAASAGRDEIVKALLGKGAQ--VNAVNGCTPLHYAASKNRHETIAVMLEGGANP 119
Db 74 TPFHTACSVGNLEVVKSLYDRPLKPNKITNQGVTCPLHLAVGKKWPEVSQFLIENGASV 133
Qy 120 DAKDHYEATAMHRAAAGNLMKMIHLL--YYKASTNIQDTGNTPLHLACDEERVEEAKLL 178
Db 134 RIKDFNQLPLHRAASVGSGLKLIELLCGLGKSAVNWDQKQGTPLFLHALAEHGDAVLL 193
Qy 179 VSQASIY--IENKEK 193
Db 194 VEKYCAEYDLYDNKAK 210

RESULT 11
A57282
ankyrin-related protein unc-44 - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999
C:Accession: A57282
R:Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakulpo
J. Cell Biol. 129, 1081-1092, 1995
A:Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Cae
A:Reference number: A57282; MUID:95263663
A:Accession: A57282
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-1786 <OTS>
A:Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
C:Genetics:
A:Gene: unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:164-192/Domain: ankyrin repeat homology <AN04>
F:358-390/Domain: ankyrin repeat homology <ANI>
F:391-423/Domain: ankyrin repeat homology <AN11>

Query Match 26.48; Score 288.5; DB 2; Length 1786;
Best Local Similarity 36.08; Pred. No. 5.5e-17;
Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
Qy 7 BELKESIL---ADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLQGVVNVNKKDDAGHSP 63
Db 537 EEVAGILLDHNADKTLTK---KGFTPLHLASKYGNLEVRLLERGPVVDIEGKNQVTP 593

Qy 64 LHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHETIAVMLEGGANPDADK 123
Db 594 LHVAHYNNDKVAMLLLENGASAKAAKNGYTPHLIAAKKNOMEIASTLLQFKADPNKAS 653

Qy 124 HYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKETISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVQAQILYNNGA 713

Qy 184 SIYIENKEEKTPLQVA--KGGIGLILKRWVE 212
Db 714 EINSKTNAGYTPPLHVACHFGQLNMV-KFLVE 743

RESULT 12
T15346
elegans ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C:Accession: T15346
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15346
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Residues: 1-1815 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1; CESP:un
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
C:Superfamily: ankyrin; ankyrin repeat homology
F:358-390/Domain: ankyrin repeat homology <ANR>

Query Match 26.48; Score 288.5; DB 2; Length 1815;
Best Local Similarity 36.08; Pred. No. 5.5e-17;
Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
Qy 7 BELKESIL---ADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLQGVVNVNKKDDAGHSP 63
Db 537 EEVAGILLDHNADKTLTK---KGFTPLHLASKYGNLEVRLLERGPVVDIEGKNQVTP 593

Qy 64 LHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHETIAVMLEGGANPDADK 123
Db 594 LHVAHYNNDKVAMLLLENGASAKAAKNGYTPHLIAAKKNOMEIASTLLQFKADPNKAS 653

Qy 124 HYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKETISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVQAQILYNNGA 713

Qy 184 SIYIENKEEKTPLQVA--KGGIGLILKRWVE 212
Db 714 EINSKTNAGYTPPLHVACHFGQLNMV-KFLVE 743

RESULT 13

T15344
ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T15344; T15345
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15344
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1867 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208874; PIDN:AAA93444.1
A:Accession: T15345
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1806, 'G', 1808, 'T' <GA2>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208876; PIDN:AAA93446.1
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1781
A:Superfamily: ankyrin; ankyrin repeat homology
F:358-390/Domain: ankyrin repeat homology <ANR>

Query Match 26.4%; Score 288.5; DB 2; Length 1867;
Best Local Similarity 36.0%; Pred. No. 5.8e-17;
Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
QY 7 EELKESIL---ADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDAGWSP 63
Db 537 EEVAGILLDHNDKTLTK---KGFTPLHLASKYGNLEVVRLLEGRTPVDIEGKNQVTP 593
QY 64 LHIAASGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHIEIAVMLLEGGANPDARD 123
Db 594 LHVAHYNNDKVAMLLLENGASAKAAKNGYTPHLHIAAKNQMEIASTLLQFKADPNKAS 653
QY 124 HYATAMHRAAAKGNLKMIIHLLYYKASTNIQDTEGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKEISGLLIENGSDVAKANNGLTAMHLCAQEDHVPVQAIIYNNGA 713
QY 184 SIYIENKEEKTPLQVA--KGLGLILKRMVE 212
Db 714 EINSKTNAGYTPLVHACHFGQLNMV-KFLVE 743

RESULT 14

T15347
ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 31-Jan-2000
C:Accession: T15347
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15347
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2039 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1; CESP:unc-44
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1799
C:Superfamily: ankyrin; ankyrin repeat homology

Query Match 26.4%; Score 288.5; DB 2; Length 2039;
Best Local Similarity 36.0%; Pred. No. 6.5e-17;

Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
QY 7 EELKESIL---ADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDAGWSP 63
Db 537 EEVAGILLDHNDKTLTK---KGFTPLHLASKYGNLEVVRLLEGRTPVDIEGKNQVTP 593
QY 64 LHIAASGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHIEIAVMLLEGGANPDARD 123
Db 594 LHVAHYNNDKVAMLLLENGASAKAAKNGYTPHLHIAAKNQMEIASTLLQFKADPNKAS 653
QY 124 HYATAMHRAAAKGNLKMIIHLLYYKASTNIQDTEGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKEISGLLIENGSDVAKANNGLTAMHLCAQEDHVPVQAIIYNNGA 713
QY 184 SIYIENKEEKTPLQVA--KGLGLILKRMVE 212
Db 714 EINSKTNAGYTPLVHACHFGQLNMV-KFLVE 743

RESULT 15

S37771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: S37771
R:Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found
A:Reference number: S37771; MUID:93252825
A:Accession: S37771
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1848 <BIR>
A:Cross-references: EMBL:X6063; NID:g311816; PIDN:CAA48801.1; PID:g311817
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:48-80/Domain: ankyrin repeat homology <AN01>
F:81-113/Domain: ankyrin repeat homology <AN02>
F:114-146/Domain: ankyrin repeat homology <AN03>
F:147-175/Domain: ankyrin repeat homology <AN04>
F:176-208/Domain: ankyrin repeat homology <AN05>
F:209-241/Domain: ankyrin repeat homology <AN06>
F:242-274/Domain: ankyrin repeat homology <AN07>
F:275-307/Domain: ankyrin repeat homology <AN08>
F:308-340/Domain: ankyrin repeat homology <AN09>
F:341-373/Domain: ankyrin repeat homology <AN10>
F:374-406/Domain: ankyrin repeat homology <AN11>
F:407-439/Domain: ankyrin repeat homology <AN12>
F:440-472/Domain: ankyrin repeat homology <AN13>
F:473-505/Domain: ankyrin repeat homology <AN14>
F:506-538/Domain: ankyrin repeat homology <AN15>
F:539-571/Domain: ankyrin repeat homology <AN16>
F:572-604/Domain: ankyrin repeat homology <AN17>
F:605-637/Domain: ankyrin repeat homology <AN18>
F:638-670/Domain: ankyrin repeat homology <AN19>
F:671-703/Domain: ankyrin repeat homology <AN20>
F:704-736/Domain: ankyrin repeat homology <AN21>
F:737-769/Domain: ankyrin repeat homology <AN22>
F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match 26.3%; Score 287.5; DB 2; Length 1848;
Best Local Similarity 32.9%; Pred. No. 7e-17;
Matches 78; Conservative 31; Mismatches 89; Indels 39; Gaps 5;

QY 13 ILADK---SLATTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDAGWS-----62
Db 261 LLLDRGAQIETRT-KDETLPLHCAARNHVRSEILLDHGAFIOAKTKGLSPFIHMAAG 319
QY 63 -----PLHIAASAGRDEIVKALLGKGAQVNAVNGCTPL 97
Db 320 DHLDCVRLLLQYNAEIDITLDHLTPLHVAACHGHRHVRVAKVLLDKGAKPNSRALNGFTPL 379

	98	HYAAKKNHRELAVLLEGGANPDADKHYEATAMHRAAAGNLKMTHILLYYKASTNIQDT	157
Qy		: : : : : : : : : : : : : : : :	
	380	HIACKNHHIRWELLUKTASIDAVTEGULPLUHVASFVCHLPVLVKNLQRGASPNSNV	439
Db		: : : : : : : : : : : : : : : :	
	158	EGNTPLHLACDEERVEBAKLIVSQGASIYIENKEETPLQVAK--GGGLILIKRWVE	212
Qy		: : : : : : : : : : : : : : : :	
	440	KVETPLHMAARAGTEVAYLLQNKKANAKAKDDQTPLRCAARTIGHTGMV- KLILE	495
Db		: : : : : : : : : : : : : : : :	

Search completed: August 13, 2001, 07:46:42
Job time: 139 sec

THIS PAGE BLANK (USPTO)

Result No.	Query	Score	Match	Length	DB	ID	Description
1	294	26.9	3924	1	ANK2_HUMAN	Q01484	homo sapien
2	290.5	26.6	328	1	YGAX_YEAST	P50086	saccharomyc
3	287.5	26.3	1862	1	ANK1_MOUSE	Q02357	mus musculus
4	286.5	26.2	1880	1	ANK1_HUMAN	P16157	homo sapien
5	276.5	25.0	234	1	YD57_SCHPO	P10311	schizosacch
6	247.5	22.6	1401	1	LATA_LATWA	P23631	latrodectus
7	247	22.6	1431	1	DAPK_HUMAN	P53355	homo sapien
8	245	22.4	832	1	ANR3_HUMAN	P57078	homo sapien
9	237	21.7	752	1	PA26_MOUSE	P97819	mus musculus
10	236.5	21.6	439	1	AKR_ARATH	Q05753	arabidopsis
11	232	21.2	751	1	PA36_RAT	P97570	rattus norv
12	225	20.6	323	1	ANKH_CHRV1	Q06527	chromatium
13	224.5	20.5	898	1	KBF2_HUMAN	Q00653	homo sapien
14	222.5	20.4	806	1	PA36_HUMAN	Q06733	homo sapien
15	217.5	19.9	452	1	ILK_MOUSE	O55222	mus musculus
16	215	19.7	2703	1	NOTC_DROME	P07207	drosophila
17	214	19.6	2318	1	NTC3_MOUSE	Q61982	mus musculus
18	212.5	19.4	451	1	ILK_CAVPO	P57044	cavia porce
19	211.5	19.4	452	1	ILK1_HUMAN	Q13418	homo sapien
20	211.5	19.4	452	1	ILK2_HUMAN	P57043	homo sapien
21	210	19.2	656	1	FEM1_CAEEL	P17221	caenorhabdi
22	208.5	19.1	1964	1	NTC4_MOUSE	P31695	mus musculus
23	207	18.9	2531	1	NTC1_RAT	Q07008	rattus norv
24	204.5	18.7	2444	1	NTC1_HUMAN	PA6531	homo sapien
25	202	18.5	679	1	RNSA_MOUSE	Q05921	mus musculus
26	201	18.4	446	1	BCL3_HUMAN	P20749	homo sapien
27	197.5	18.1	2437	1	NOTC_BRARE	PA6530	brachydanio
28	194	17.7	741	1	RNSA_HUMAN	Q05823	homo sapien
29	194	17.7	764	1	AKF1_YEAST	P39010	saccharomyc
30	194	17.7	2524	1	NOTC_XENLA	P21783	xenopus lae
31	193.5	17.7	984	1	NOTC_CHICK	Q04861	gallus gall
32	189.5	17.3	347	1	GABC_HUMAN	Q06545	homo sapien
33	189.5	17.3	383	1	GABB_HUMAN	Q06547	homo sapien

CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z26634; CAB42644.1; -
CC EMBL; X56957; CAA40278.1; -
CC EMBL; X56958; CAA40279.2; -
CC EMBL; M37123; AAA62828.1; -
CC PIR; S14533; S14533.
CC PIR; A39643; A39643.
CC PIR; B39643; B39643.
CC PIR; S14569; S14569.
CC HSP; Q00420; IAWC.
CC MIM; 106410; -
CC InterPro; IPR000488; -
CC InterPro; IPR000906; -
CC InterPro; IPR002110; -
CC Pfam; PF00791; ZUS; 1.
CC Pfam; PF0023; ank; 22.
CC Pfam; PF00531; death; 1.
CC PROSITE; PS50088; ANK_REPEAT; 20.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC PROSITE; PS50017; DEATH_DOMAIN; 1.
CC Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
CC Phosphorylation; Multigene family.
CC REPEAT 63 92 ANK 1.
CC REPEAT 96 125 ANK 2.
CC REPEAT 129 158 ANK 3.
CC REPEAT 162 191 ANK 4.
CC REPEAT 193 220 ANK 5.
CC REPEAT 232 261 ANK 6.
CC REPEAT 265 294 ANK 7.
CC REPEAT 298 327 ANK 8.
CC REPEAT 331 360 ANK 9.
CC REPEAT 364 393 ANK 10.
CC REPEAT 397 426 ANK 11.
CC REPEAT 430 459 ANK 12.
CC REPEAT 463 492 ANK 13.
CC REPEAT 496 525 ANK 14.
CC REPEAT 529 558 ANK 15.
CC REPEAT 562 591 ANK 16.
CC REPEAT 595 624 ANK 17.
CC REPEAT 628 657 ANK 18.
CC REPEAT 661 690 ANK 19.
CC REPEAT 694 723 ANK 20.
CC REPEAT 727 756 ANK 21.
CC REPEAT 760 789 ANK 22.
CC REPEAT 793 822 ANK 23.
CC REPEAT 825 854 ANK 24.
CC REPEAT 857 886 ANK 25.
CC REPEAT 889 918 ANK 26.
CC REPEAT 921 950 ANK 27.
CC REPEAT 953 982 ANK 28.
CC REPEAT 985 1014 ANK 29.
CC REPEAT 1017 1046 ANK 30.
CC REPEAT 1049 1078 ANK 31.
CC REPEAT 1081 1110 ANK 32.
CC REPEAT 1113 1142 ANK 33.
CC REPEAT 1145 1174 ANK 34.
CC REPEAT 1177 1206 ANK 35.
CC REPEAT 1209 1238 ANK 36.
CC REPEAT 1241 1270 ANK 37.
CC REPEAT 1273 1302 ANK 38.
CC REPEAT 1305 1334 ANK 39.
CC REPEAT 1337 1366 ANK 40.
CC REPEAT 1369 1398 ANK 41.
CC REPEAT 1401 1430 ANK 42.
CC REPEAT 1433 1462 ANK 43.
CC REPEAT 1465 1494 ANK 44.
CC REPEAT 1497 1526 ANK 45.
CC REPEAT 1529 1558 ANK 46.
CC REPEAT 1561 1590 ANK 47.
CC REPEAT 1593 1622 ANK 48.
CC REPEAT 1625 1654 ANK 49.
CC REPEAT 1657 1686 ANK 50.
CC REPEAT 1689 1718 ANK 51.
CC REPEAT 1721 1750 ANK 52.
CC REPEAT 1753 1782 ANK 53.
CC REPEAT 1785 1814 ANK 54.
CC REPEAT 1817 1846 ANK 55.
CC REPEAT 1849 1879 ANK 56.
CC REPEAT 1881 1910 ANK 57.
CC REPEAT 1913 1942 ANK 58.
CC REPEAT 1945 1976 ANK 59.
CC REPEAT 1977 2008 ANK 60.
CC REPEAT 2009 2039 ANK 61.
CC REPEAT 2041 2072 ANK 62.
CC REPEAT 2073 2104 ANK 63.
CC REPEAT 2105 2136 ANK 64.
CC REPEAT 2137 2168 ANK 65.
CC REPEAT 2169 2200 ANK 66.
CC REPEAT 2201 2232 ANK 67.
CC REPEAT 2233 2264 ANK 68.
CC REPEAT 2265 2296 ANK 69.
CC REPEAT 2297 2328 ANK 70.
CC REPEAT 2329 2360 ANK 71.
CC REPEAT 2361 2392 ANK 72.
CC REPEAT 2393 2424 ANK 73.
CC REPEAT 2425 2456 ANK 74.
CC REPEAT 2457 2488 ANK 75.
CC REPEAT 2489 2520 ANK 76.
CC REPEAT 2521 2552 ANK 77.
CC REPEAT 2553 2584 ANK 78.
CC REPEAT 2585 2616 ANK 79.
CC REPEAT 2617 2648 ANK 80.
CC REPEAT 2649 2680 ANK 81.
CC REPEAT 2681 2712 ANK 82.
CC REPEAT 2713 2744 ANK 83.
CC REPEAT 2745 2776 ANK 84.
CC REPEAT 2777 2808 ANK 85.
CC REPEAT 2809 2840 ANK 86.
CC REPEAT 2841 2872 ANK 87.
CC REPEAT 2873 2904 ANK 88.
CC REPEAT 2905 2936 ANK 89.
CC REPEAT 2937 2968 ANK 90.
CC REPEAT 2969 3000 ANK 91.
CC REPEAT 3001 3032 ANK 92.
CC REPEAT 3033 3064 ANK 93.
CC REPEAT 3065 3096 ANK 94.
CC REPEAT 3097 3128 ANK 95.
CC REPEAT 3129 3160 ANK 96.
CC REPEAT 3161 3192 ANK 97.
CC REPEAT 3193 3224 ANK 98.
CC REPEAT 3225 3256 ANK 99.
CC REPEAT 3257 3288 ANK 100.
CC REPEAT 3289 3320 ANK 101.
CC REPEAT 3321 3352 ANK 102.
CC REPEAT 3353 3384 ANK 103.
CC REPEAT 3385 3416 ANK 104.
CC REPEAT 3417 3448 ANK 105.
CC REPEAT 3449 3480 ANK 106.
CC REPEAT 3481 3512 ANK 107.
CC REPEAT 3513 3544 ANK 108.
CC REPEAT 3545 3576 ANK 109.
CC REPEAT 3577 3608 ANK 110.
CC REPEAT 3609 3640 ANK 111.
CC REPEAT 3641 3672 ANK 112.
CC REPEAT 3673 3704 ANK 113.
CC REPEAT 3705 3736 ANK 114.
CC REPEAT 3737 3768 ANK 115.
CC REPEAT 3769 3800 ANK 116.
CC REPEAT 3801 3832 ANK 117.
CC REPEAT 3833 3864 ANK 118.
CC REPEAT 3865 3896 ANK 119.
CC REPEAT 3897 3928 ANK 120.
CC REPEAT 3929 3960 ANK 121.
CC REPEAT 3961 3992 ANK 122.
CC REPEAT 3993 4024 ANK 123.
CC REPEAT 4025 4056 ANK 124.
CC REPEAT 4057 4088 ANK 125.
CC REPEAT 4089 4120 ANK 126.
CC REPEAT 4121 4152 ANK 127.
CC REPEAT 4153 4184 ANK 128.
CC REPEAT 4185 4216 ANK 129.
CC REPEAT 4217 4248 ANK 130.
CC REPEAT 4249 4280 ANK 131.
CC REPEAT 4281 4312 ANK 132.
CC REPEAT 4313 4344 ANK 133.
CC REPEAT 4345 4376 ANK 134.
CC REPEAT 4377 4408 ANK 135.
CC REPEAT 4409 4440 ANK 136.
CC REPEAT 4441 4472 ANK 137.
CC REPEAT 4473 4504 ANK 138.
CC REPEAT 4505 4536 ANK 139.
CC REPEAT 4537 4568 ANK 140.
CC REPEAT 4569 4600 ANK 141.
CC REPEAT 4601 4632 ANK 142.
CC REPEAT 4633 4664 ANK 143.
CC REPEAT 4665 4696 ANK 144.
CC REPEAT 4697 4728 ANK 145.
CC REPEAT 4729 4760 ANK 146.
CC REPEAT 4761 4792 ANK 147.
CC REPEAT 4793 4824 ANK 148.
CC REPEAT 4825 4856 ANK 149.
CC REPEAT 4857 4888 ANK 150.
CC REPEAT 4889 4920 ANK 151.
CC REPEAT 4921 4952 ANK 152.
CC REPEAT 4953 4984 ANK 153.
CC REPEAT 4985 5016 ANK 154.
CC REPEAT 5017 5048 ANK 155.
CC REPEAT 5049 5080 ANK 156.
CC REPEAT 5081 5112 ANK 157.
CC REPEAT 5113 5144 ANK 158.
CC REPEAT 5145 5176 ANK 159.
CC REPEAT 5177 5208 ANK 160.
CC REPEAT 5209 5240 ANK 161.
CC REPEAT 5241 5272 ANK 162.
CC REPEAT 5273 5304 ANK 163.
CC REPEAT 5305 5336 ANK 164.
CC REPEAT 5337 5368 ANK 165.
CC REPEAT 5369 5400 ANK 166.
CC REPEAT 5401 5432 ANK 167.
CC REPEAT 5433 5464 ANK 168.
CC REPEAT 5465 5496 ANK 169.
CC REPEAT 5497 5528 ANK 170.
CC REPEAT 5529 5560 ANK 171.
CC REPEAT 5561 5592 ANK 172.
CC REPEAT 5593 5624 ANK 173.
CC REPEAT 5625 5656 ANK 174.
CC REPEAT 5657 5688 ANK 175.
CC REPEAT 5689 5720 ANK 176.
CC REPEAT 5721 5752 ANK 177.
CC REPEAT 5753 5784 ANK 178.
CC REPEAT 5785 5816 ANK 179.
CC REPEAT 5817 5848 ANK 180.
CC REPEAT 5849 5880 ANK 181.
CC REPEAT 5881 5912 ANK 182.
CC REPEAT 5913 5944 ANK 183.
CC REPEAT 5945 5976 ANK 184.
CC REPEAT 5977 6008 ANK 185.
CC REPEAT 6009 6040 ANK 186.
CC REPEAT 6041 6072 ANK 187.
CC REPEAT 6073 6104 ANK 188.
CC REPEAT 6105 6136 ANK 189.
CC REPEAT 6137 6168 ANK 190.
CC REPEAT 6169 6200 ANK 191.
CC REPEAT 6201 6232 ANK 192.
CC REPEAT 6233 6264 ANK 193.
CC REPEAT 6265 6296 ANK 194.
CC REPEAT 6297 6328 ANK 195.
CC REPEAT 6329 6360 ANK 196.
CC REPEAT 6361 6392 ANK 197.
CC REPEAT 6393 6424 ANK 198.
CC REPEAT 6425 6456 ANK 199.
CC REPEAT 6457 6488 ANK 200.
CC REPEAT 6489 6520 ANK 201.
CC REPEAT 6521 6552 ANK 202.
CC REPEAT 6553 6584 ANK 203.
CC REPEAT 6585 6616 ANK 204.
CC REPEAT 6617 6648 ANK 205.
CC REPEAT 6649 6680 ANK 206.
CC REPEAT 6681 6712 ANK 207.
CC REPEAT 6713 6744 ANK 208.
CC REPEAT 6745 6776 ANK 209.
CC REPEAT 6777 6808 ANK 210.
CC REPEAT 6809 6840 ANK 211.
CC REPEAT 6841 6872 ANK 212.
CC REPEAT 6873 6904 ANK 213.
CC REPEAT 6905 6936 ANK 214.
CC REPEAT 6937 6968 ANK 215.
CC REPEAT 6969 7000 ANK 216.
CC REPEAT 7001 7032 ANK 217.
CC REPEAT 7033 7064 ANK 218.
CC REPEAT 7065 7096 ANK 219.
CC REPEAT 7097 7128 ANK 220.
CC REPEAT 7129 7160 ANK 221.
CC REPEAT 7161 7192 ANK 222.
CC REPEAT 7193 7224 ANK 223.
CC REPEAT 7225 7256 ANK 224.
CC REPEAT 7257 7288 ANK 225.
CC REPEAT 7289 7320 ANK 226.
CC REPEAT 7321 7352 ANK 227.
CC REPEAT 7353 7384 ANK 228.
CC REPEAT 7385 7416 ANK 229.
CC REPEAT 7417 7448 ANK 230.
CC REPEAT 7449 7480 ANK 231.
CC REPEAT 7481 7512 ANK 232.
CC REPEAT 7513 7544 ANK 233.
CC REPEAT 7545 7576 ANK 234.
CC REPEAT 7577 7608 ANK 235.
CC REPEAT 7609 7640 ANK 236.
CC REPEAT 7641 7672 ANK 237.
CC REPEAT 7673 7704 ANK 238.
CC REPEAT 7705 7736 ANK 239.
CC REPEAT 7737 7768 ANK 240.
CC REPEAT 7769 7800 ANK 241.
CC REPEAT 7801 7832 ANK 242.
CC REPEAT 7833 7864 ANK 243.
CC REPEAT 7865 7896 ANK 244.
CC REPEAT 7897 7928 ANK 245.
CC REPEAT 7929 7960 ANK 246.
CC REPEAT 7961 7992 ANK 247.
CC REPEAT 7993 8024 ANK 248.
CC REPEAT 8025 8056 ANK 249.
CC REPEAT 8057 8088 ANK 250.
CC REPEAT 8089 8120 ANK 251.
CC REPEAT 8121 8152 ANK 252.
CC REPEAT 8153 8184 ANK 253.
CC REPEAT 8185 8216 ANK 254.
CC REPEAT 8217 8248 ANK 255.
CC REPEAT 8249 8280 ANK 256.
CC REPEAT 8281 8312 ANK 257.
CC REPEAT 8313 8344 ANK 258.
CC REPEAT 8345 8376 ANK 259.
CC REPEAT 8377 8408 ANK 260.
CC REPEAT 8409 8440 ANK 261.
CC REPEAT 8441 8472 ANK 262.
CC REPEAT 8473 8504 ANK 263.
CC REPEAT 8505 8536 ANK 264.
CC REPEAT 8537 8568 ANK 265.
CC REPEAT 8569 8600 ANK 266.
CC REPEAT 8601 8632 ANK 267.
CC REPEAT 8633 8664 ANK 268.
CC REPEAT 8665 8696 ANK 269.
CC REPEAT 8697 8728 ANK 270.
CC REPEAT 8729 8760 ANK 271.
CC REPEAT 8761 8792 ANK 272.
CC REPEAT 8793 8824 ANK 273.
CC REPEAT 8825 8856 ANK 274.
CC REPEAT 8857 8888 ANK 275.
CC REPEAT 8889 8920 ANK 276.
CC REPEAT 8921 8952 ANK 277.
CC REPEAT 8953 8984 ANK 278.
CC REPEAT 8985 9016 ANK 279.
CC REPEAT 9017 9048 ANK 280.
CC REPEAT 9049 9080 ANK 281.
CC REPEAT 9081 9112 ANK 282.
CC REPEAT 9113 9144 ANK 283.
CC REPEAT 9145 9176 ANK 284.
CC REPEAT 9177 9208 ANK 285.
CC REPEAT 9209 9240 ANK 286.
CC REPEAT 9241 9272 ANK 287.
CC REPEAT 9273 9304 ANK 288.
CC REPEAT 9305 9336 ANK 289.
CC REPEAT 9337 9368 ANK 290.
CC REPEAT 9369 9400 ANK 291.
CC REPEAT 9401 9432 ANK 292.
CC REPEAT 9433 9464 ANK 293.
CC REPEAT 9465 9496 ANK 294.
CC REPEAT 9497 9528 ANK 295.
CC REPEAT 9529 9560 ANK 296.
CC REPEAT 9561 9592 ANK 297.
CC REPEAT 9593 9624 ANK 298.
CC REPEAT 9625 9656 ANK 299.
CC REPEAT 9657 9688 ANK 300.
CC REPEAT 9689 9720 ANK 301.
CC REPEAT 9721 9752 ANK 302.
CC REPEAT 9753 9784 ANK 303.
CC REPEAT 9785 9816 ANK 304.
CC REPEAT 9817 9848 ANK 305.
CC REPEAT 9849 9880 ANK 306.
CC REPEAT 9881 9912 ANK 307.
CC REPEAT 9913 9944 ANK 308.
CC REPEAT 9945 9976 ANK 309.
CC REPEAT 9977 10008 ANK 310.
CC REPEAT 10009 10040 ANK 311.
CC REPEAT 10041 10072 ANK 312.
CC REPEAT 10073 10104 ANK 313.
CC REPEAT 10105 10136 ANK 314.
CC REPEAT 10137 10168 ANK 315.
CC REPEAT 10169 10200 ANK 316.
CC REPEAT 10201 10232 ANK 317.
CC REPEAT 10233 10264 ANK 318.
CC REPEAT 10265 10296 ANK 319.
CC REPEAT 10297 10328 ANK 320.
CC REPEAT 10329 10360 ANK 321.
CC REPEAT 10361 10392 ANK 322.
CC REPEAT 10393 10424 ANK 323.
CC REPEAT 10425 10456 ANK 324.
CC REPEAT 10457 10488 ANK 325.
CC REPEAT 10489 10520 ANK 326.
CC REPEAT 10521 10552 ANK 327.
CC REPEAT 10553 10584 ANK 328.
CC REPEAT 10585 10616 ANK 329.
CC REPEAT 10617 10648 ANK 330.
CC REPEAT 10649 10680 ANK 331.
CC REPEAT 10681 10712 ANK 332.
CC REPEAT 10713 10744 ANK 333.
CC REPEAT 10745 10776 ANK 334.
CC REPEAT 10777 10808 ANK 335.
CC REPEAT 10809 10840 ANK 336.
CC REPEAT 10841 10872 ANK 337.
CC REPEAT 10873 10904 ANK 338.
CC REPEAT 10905 10936 ANK 339.
CC REPEAT 10937 10968 ANK 340.
CC REPEAT 10969 11000 ANK 341.
CC REPEAT 11001 11032 ANK 342.
CC REPEAT 11033 11064 ANK 343.
CC REPEAT 11065 11096 ANK 344.
CC REPEAT 11097 11128 ANK 345.
CC REPEAT 11129 11160 ANK 346.
CC REPEAT 11161 11192 ANK 347.
CC REPEAT 11193 11224 ANK 348.
CC REPEAT 11225 11256 ANK 349.
CC REPEAT 11257 11288 ANK 350.
CC REPEAT 11289 11320 ANK 351.
CC REPEAT 11321 11352 ANK 352.
CC REPEAT 11353 11384 ANK 353.
CC REPEAT 11385 11416 ANK 354.
CC REPEAT 11417 11448 ANK 355.
CC REPEAT 11449 11480 ANK 356.
CC REPEAT 11481 11512 ANK 357.
CC REPEAT 11513 11544 ANK 358.
CC REPEAT 11545 11576 ANK 359.
CC REPEAT 11577 11608 ANK 360.
CC REPEAT 11609 11640 ANK 361.
CC REPEAT 11641 11672 ANK 362.
CC REPEAT 11673 11704 ANK 363.
CC REPEAT 11705 11736 ANK 364.
CC REPEAT 11737 11768 ANK 365.
CC REPEAT 11769 11800 ANK 366.
CC REPEAT 11801 11832 ANK 367.
CC REPEAT 11833 11864 ANK 368.
CC REPEAT 11865 11896 ANK 369.
CC REPEAT 11897 11928 ANK 370.
CC REPEAT 11929 11960 ANK 371.
CC REPEAT 11961 11992 ANK 372.
CC REPEAT 11993 12024 ANK 373.
CC REPEAT 12025 12056 ANK 374.
CC REPEAT 12057 12088 ANK 375.
CC REPEAT 12089 12120 ANK 376.
CC REPEAT 12121 12152 ANK 377.
CC REPEAT 12153 12184 ANK 378.
CC REPEAT 12185 12216 ANK 379.
CC REPEAT 12217 12248 ANK 380.
CC REPEAT 12249 12280 ANK 381.
CC REPEAT 12281 12312 ANK 382.
CC REPEAT 12313 12344 ANK 383.
CC REPEAT 12345 12376 ANK 384.
CC REPEAT 12377 12408 ANK 385.
CC REPEAT 12409 12440 ANK 386.
CC REPEAT 12441 12472 ANK 387.
CC REPEAT 12473 12504 ANK 388.
CC REPEAT 12505 12536 ANK 389.
CC REPEAT 12537 12568 ANK 390.
CC REPEAT 12569 12600 ANK 391.
CC REPEAT 12601 12632 ANK 392.
CC REPEAT 12633 12664 ANK 393.
CC REPEAT 12665 12696 ANK 394.
CC REPEAT 12697 12728 ANK 395.
CC REPEAT 12729 12760 ANK 396.
CC REPEAT 12761 12792 ANK 397.
CC REPEAT 12793 12824 ANK 398.
CC REPEAT 12825 12856 ANK 399.
CC REPEAT 12857 12888 ANK 400.
CC REPEAT 12889 12920 ANK 401.
CC REPEAT 12921 12952 ANK 402.
CC REPEAT 12953 12984 ANK 403.
CC REPEAT 12985 13016 ANK 404.
CC REPEAT 13017 13048 ANK 405.
CC REPEAT 13049 13080 ANK 406.
CC REPEAT 13081 13112 ANK 407.
CC REPEAT 13113 13144 ANK 408.
CC REPEAT 13145 13176 ANK 409.
CC REPEAT 13177 13208 ANK 410.
CC REPEAT 13209 13240 ANK 411.
CC REPEAT 13241 13272 ANK 412.
CC REPEAT 13273 13304 ANK 413.
CC REPEAT 13305 13336 ANK 414.
CC REPEAT 13337 13368 ANK 415.
CC REPEAT 13369 13400 ANK 416.
CC REPEAT 13401 13432 ANK 417.
CC REPEAT 13433 13464 ANK 418.
CC REPEAT 13465 13496 ANK 419.
CC REPEAT 13497 13528 ANK 420.
CC REPEAT 13529 13560 ANK 421.
CC REPEAT 13561 13592 ANK 422.
CC REPEAT 13593 13624 ANK 423.
CC REPEAT 13625 13656 ANK 424.
CC REPEAT 13657 13688 ANK 425.
CC REPEAT 13689 13720 ANK 426.
CC REPEAT 13721 13752 ANK 427.
CC REPEAT 13753 13784 ANK 428.
CC REPEAT 13785 13816 ANK 429.
CC REPEAT 13817 13848 ANK 430.
CC REPEAT 13849 13880 ANK 431.
CC REPEAT 13881 13912 ANK 432.
CC REPEAT 13913 13944 ANK 433.
CC REPEAT 13945 13976 ANK 434.
CC REPEAT 13977 14008 ANK 435.
CC REPEAT 14009 14040 ANK 436.
CC REPEAT 14041 14072 ANK 437.
CC REPEAT 14073 14104 ANK 438.
CC REPEAT 14105 14136 ANK 439.
CC REPEAT 14137 14168 ANK 440.
CC REPEAT 14169 14200 ANK 441.
CC REPEAT 14201 14232 ANK 442.
CC REPEAT 14233 14264 ANK 443.
CC REPEAT 14265 14296 ANK 444.
CC REPEAT 14297 14328 ANK 445.
CC REPEAT 14329 14360 ANK 446.
CC REPEAT 14361 14392 ANK 447.
CC REPEAT 14393 14424 ANK 448.
CC REPEAT 14425 14456 ANK 449.
CC REPEAT 14457 14488 ANK 450.
CC REPEAT 14489 14520 ANK 451.
CC REPEAT 14521 14552 ANK 452.
CC REPEAT 14553 14584 ANK 453.
CC REPEAT 14585 14616 ANK 454.
CC REPEAT 14617 14648 ANK 455.
CC REPEAT 14649 14680 ANK 456.
CC REPEAT 14681 14712 ANK 457.
CC REPEAT 14713 14744 ANK 458.
CC REPEAT 14745 14776 ANK 459.
CC REPEAT 14777 14808 ANK 460.
CC REPEAT 14809 14840 ANK 461.
CC REPEAT 14841 14872 ANK 462.
CC REPEAT 14873 14904 ANK 463.
CC REPEAT 14905 14936 ANK 464.
CC REPEAT 14937 14968 ANK 465.
CC REPEAT 14969 15000 ANK 466.
CC REPEAT 15001 15032 ANK 467.
CC REPEAT 15033 15064 ANK 468.
CC REPEAT 15065 15096 ANK 469.
CC REPEAT 15097 15128 ANK 470.
CC REPEAT 15129 15160 ANK 471.
CC REPEAT 15161 15192 ANK 472.
CC REPEAT 15193 15224 ANK 473.
CC REPEAT 15225 15256 ANK 474.
CC REPEAT 15257 15288 ANK 475.
CC REPEAT 15289 15320 ANK 476.
CC REPEAT 15321 15352 ANK 477.
CC REPEAT 15353 15384 ANK 478.
CC REPEAT 15385 15416 ANK 479.
CC REPEAT 15417 15448 ANK 480.
CC REPEAT 15449 15480 ANK 481.
CC REPEAT 15481 15512 ANK 482.
CC REPEAT 15513 15544 ANK 483.
CC REPEAT 15545 15576 ANK 484.
CC REPEAT 15577 15608 ANK 485.
CC REPEAT 15609 15640 ANK 486.
CC REPEAT 15641 15672 ANK 487.
CC REPEAT 15673 15704 ANK 488.
CC REPEAT 15705 15736 ANK 489.
CC REPEAT 15737 15768 ANK 490.
CC REPEAT 15769 15800 ANK 491.
CC REPEAT 15801 15832 ANK 492.
CC REPEAT 15833 15864 ANK 493.
CC REPEAT 15865 15896 ANK 494.
CC REPEAT 15897 15928 ANK 495.
CC REPEAT 15929 15960 ANK 496.
CC REPEAT 15961 15992 ANK 497.
CC REPEAT 15993 16024 ANK 498.
CC REPEAT 16025 16056 ANK 499.
CC REPEAT 16057 16088 ANK 500.
CC REPEAT 16089 16120 ANK 501.
CC REPEAT 16121 16152 ANK 502.
CC REPEAT 16153 16184 ANK 503.
CC REPEAT 16185 16216 ANK 504.
CC REPEAT 16217 16248 ANK 505.
CC REPEAT 16249 16280 ANK 506.
CC REPEAT 16281 16312 ANK 507.
CC REPEAT 16313 16344 ANK 508.
CC REPEAT 16345 16376 ANK 509.
CC REPEAT 16377 16408 ANK 510.
CC REPEAT 16409 16440 ANK 511.
CC REPEAT 16441 16472 ANK 512.
CC REPEAT 16473 16504 ANK 513.
CC REPEAT 16505 16536 ANK 514.
CC REPEAT 16537 16568 ANK 515.
CC REPEAT 16569 16600 ANK 516.
CC REPEAT 16601 16632 ANK 517.
CC REPEAT

FT REPEAT 139 168 ANK 5.
FT REPEAT 173 203 ANK 6.
SQ SEQUENCE 228 AA; 25616 MW; 6B30CDD034A6289 CRC64;

Query Match 26.6%; Score 290.5; DB 1; Length 228;
Best Local Similarity 38.6%; Pred. No. 4.7e-19;
Matches 76; Conservative 30; Mismatches 80; Indels 11; Gaps 5;

QY 5 KLEELKESILADKSLATRTDQSRALHWACSGAGTEIVEFLQLGVPVW---DKDDAGW 61
DQ 17 KYOELLHS---KPSLLQKDDQGRPLHWSVSFOAHEITSFLLSKMENVNLDYDDSGW 73
QY 62 SPLHIAASAGRDEIVKALIGKCAQ--VNAVNGCTPLHYAASKNRHETAVMLLEGGANP 119
DQ 74 TFPHTACSVGNLEVKSLYDRPLKPLDNKITNQGVYTLHLAVGKKWFVSQFLIENGASV 133
QY 120 DAKDHYEATAMHRAAAGNKLMIHILL--YYKASTNIQDTEGNTPLHLACDEERVEEAKLL 178
DQ 134 RIKDKFNQIPLHRAASVGSKLIELLGLGKSAVNWQDKGWTPLFHALLAEGHDAAVLL 193
QY 179 VSQASIV--IENKEEK 193
DQ 194 VEKYGAEDYLDVNDKGA 210

RESULT 3
ANK1_MOUSE
ID ANK1_MOUSE STANDARD; PRT; 1862 AA.
AC Q02357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte;
RA MEDLINE=92345717; PubMed=1386265;
RX White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RL Mamm. Genome 3:281-285(1992).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2. TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC -!- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M84756; AAA37236.1; -;
CC HSSP; Q00420; IAWC.
CC MGD; MGI:88024; Ank1.
CC InterPro; IPR000488; -;
CC InterPro; IPR000906; -;

DR InterPro; IPR002110; -;
DR Pfam; PF00791; ZU5; 1.
DR Pfam; PF00023; ank; 23.
DR Pfam; PF00531; death; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
FT DOMAIN 1 827 89 KDA DOMAIN (ANION EXCHANGE PROTEIN
FT DOMAIN 828 1386 BINDING DOMAIN).
FT DOMAIN 1387 1862 62 KDA DOMAIN (SPECTRIN BINDING
FT REPEAT 40 69 55 KDA REGULATORY DOMAIN (REGULATES
FT REPEAT 73 102 THE BINDING OF ANKYRIN TO SPECTRIN
FT REPEAT 106 135 AND THE BAND 3 PROTEIN).
FT REPEAT 139 168 ANK 1.
FT REPEAT 170 197 ANK 2.
FT REPEAT 201 230 ANK 3.
FT REPEAT 234 263 ANK 4.
FT REPEAT 267 296 ANK 5.
FT REPEAT 300 329 ANK 6.
FT REPEAT 333 362 ANK 7.
FT REPEAT 366 395 ANK 8.
FT REPEAT 399 428 ANK 9.
FT REPEAT 432 461 ANK 10.
FT REPEAT 465 494 ANK 11.
FT REPEAT 498 527 ANK 12.
FT REPEAT 531 560 ANK 13.
FT REPEAT 564 593 ANK 14.
FT REPEAT 597 626 ANK 15.
FT REPEAT 630 659 ANK 16.
FT REPEAT 663 692 ANK 17.
FT REPEAT 696 725 ANK 18.
FT REPEAT 729 758 ANK 19.
FT REPEAT 762 791 ANK 20.
FT DOMAIN 1399 1483 ANK 21.
SQ SEQUENCE 1862 AA; 204242 MW; AE6B85B5B29001E5 CRC64;

Query Match 26.3%; Score 287.5; DB 1; Length 1862;
Best Local Similarity 32.9%; Pred. No. 1.1e-17;
Matches 78; Conservative 31; Mismatches 89; Indels 39; Gaps 5;

QY 13 ILADK--SLATRTDQSRALHWACSGAGTEIVEFLQLGVPVNDKDDAGWS-----62
DQ 253 LLLDRGAQIETRT-KDELTPHCAARNGHVRISILLDHGAPIQAKTKNGLSPIHMAAQG 311
QY 63 -----PLHIAASAGRDEIVKALIGKCAQVNAVNGCTPL 97
DQ 312 DHLDCVRLLLQYNABEIDITLHDLTPHVAACHGHRVAKVLLDRGAKPNSALAGFTPL 371
QY 98 HYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAAGNKLMIHILLIYYKASTNIQDT 157
DQ 372 HIACKNHIHVRNELLTKTGASIDAVTESGLTPLHVASFMGHLPTVKNLLQRCASPNVSNV 431
QY 158 EGNTPHLHACDEERVEEAKLLVSQASIVYIENKEEKTPLQVAK--GGGLGIUKRWE 212
DQ 432 KVETPLHMAARAGHTEVAKYLLQNKAKANAKAKDDQTPHCAARIGHTGMV-KLLLE 487

RESULT 4
ANK1_HUMAN
ID ANK1_HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).
GN ANK1 OR ANK.
OS Homo sapiens (Human).

Db 792 KTSNGATPLHLATFKGSOAALLLNNEVNRDTEGOMPIGHAAMTGLDVAQAIIIS 851
QY 149 KAS-TNIOTEGTNPPLHLACDEEREEAKLLVSOGASIIYIENKEETPLQV--AKGGLGL 205
Db 852 DAFVVDIEDKNSDTPLNLAQAQSHIDVIRYIDGADINTRNKKGLAPLAFSKKGNLDM 911
QY 206 I 206
Db 912 V 912
RESULT 7
DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.-) (DAP KINASE 1).
GN DAPK1 OR DAPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129831; PubMed=7828849;
RA Daiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD
RT protein as potential mediators of the gamma interferon-induced cell
RT death.";
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL
CC DEATH.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76104; CAA53712.1; -
DR HSP; Q63450; IAO6.
DR MIM; 600831; -
DR InterPro; IPR000488; -
DR InterPro; IPR000719; -
DR InterPro; IPR002110; -
DR InterPro; IPR002290; -
DR Pfam; PF00023; ank. 8.
DR Pfam; PF00531; death. 1.
DR Pfam; PF00069; pkinase. 1.
DR PROSITE; PS50088; ANK_REPEAT; 6.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; Calmodulin-binding;
KW phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
FT DOMAIN 13 266 PROTEIN KINASE.
FT DOMAIN 267 334 CALMODULIN-BINDING.
FT REPEAT 378 407 ANK 1.
FT REPEAT 411 440 ANK 2.

FT REPEAT 444 473 ANK 3.
FT REPEAT 478 507 ANK 4.
FT REPEAT 511 540 ANK 5.
FT REPEAT 544 573 ANK 6.
FT REPEAT 577 606 ANK 7.
FT REPEAT 610 639 ANK 8.
FT REPEAT 639 905 ANK 9.
FT REPEAT 876 905 ANK 10.
FT REPEAT 1163 1197 ANK 10.
FT DOMAIN 1313 1397 DEATH.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1431 AA; 160017 MW; 9EB4811004A155B CRC64;
Query Match 22.6%; Score 247; DB 1; Length 1431;
Best Local Similarity 29.0%; Pred. No. 3.4e-14;
Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;
QY 24 DQDRTALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDE----- 74
Db 409 DRGGSNAVYWAARHGHDVDTLKELSENKCPDLVKKSGEMALHVAARYGHADVAQVTCRAS 468
QY 75 -----IVKALLGKGAQYNVAVNGCTPLHYAASKNRHETA 109
Db 469 AQIPISRTKEETPLHCAAWHGYYSVAKALCEAGCNVNIKNREGETPLLTASARGYHDIV 528
QY 110 VMLLEGANPDADKHYEATAMHRAAAKGLMKHILLYYKASTNIQDTEGNTPLHLACDE 169
Db 529 ECLAHGADLNACDDGHTALHLAVRCQMEVIKYLLSQGCFVDYQDRHGNTPLHVACKD 588
QY 170 ERVEEAKLVSGASIIYIENKEETPLQVA 199
Db 589 GNPVIVVLCALCEANCLDISNKYGRTPHLA 618
RESULT 8
ANR3_HUMAN STANDARD; PRT; 832 AA.
AC P57078;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE ANKR3 (EC 2.7.1.-) (ANKYRIN REPEAT
DE DOMAIN PROTEIN 3).
GN ANKR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Folley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riessmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Vaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=93300842; PubMed=8390993;
RA Dolata M.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA Cusanovich M.A.;
RT "Nucleotide sequence of the heme subunit of flavocytochrome c from
RT the purple phototrophic bacterium, Chromatium vinosum. A 2.6-kilobase
RT pair DNA fragment contains two multiheme cytochromes, a flavoprotein,
RT and a homolog of human ankryrin."
RL J. Biol. Chem. 268:14426-14431(1993).
CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13419; AAA23315.1; -
CC HSSP; Q00420; IAWC.
CC InterPro: IPR002110; -
CC Pfam: PF00023; ank; 8.
CC PROSITE; PS50088; ANK_REPEAT; 7.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC Signal; ANK repeat; Repeat.
CC SIGNAL; 1 28 POTENTIAL.
CC CHAIN 29 323 ANKYRIN HOMOLOG.
CC REPEAT 31 60 ANK 1.
CC REPEAT 64 93 ANK 2.
CC REPEAT 97 126 ANK 3.
CC REPEAT 130 159 ANK 4.
CC REPEAT 163 192 ANK 5.
CC REPEAT 196 225 ANK 6.
CC REPEAT 229 258 ANK 7.
CC REPEAT 262 291 ANK 8.
CC SEQUENCE 323 AA; 33453 MW; 3976D34A9BD607DF CRC64;
SQ
Query Match 20.6%; Score 225; DB 1; Length 323;
Best Local Similarity 29.7%; Pred. No. 5.3e-13;
Matches 65; Conservative 34; Mismatches 82; Indels 38; Gaps 2;
QY 19 LATRTDQSRFA-----LHWACSGHTEIVEFLQLGVFVNDKDGWSPLHIAASGRD 73
DQ 85 LSRGADVARTVAGCTPLTFAEAGCHIGISALLERGAHVHRTSGWDALMIAERHGIT 144
QY 74 EIVKALLGGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPDADKHVEATAMHRA 133
DQ 145 DMVEQLFKGADPKAADREGRTALMQAASKGTGVLPLIEGGADLEARDKQATALLIA 204
QY 134 A-----AKGNLKHIIILYYKASTNIQDTSGN 160
DQ 205 ADQOGAGAVETLAGIQAOLDAVLGSLTALILVAGHGVAMVERLLAMGADPNQDRKGT 264
QY 161 TPLHLACDEERVEEAKLVISQAGSIYENKEKTPLOVA 199
DQ 265 TALMEAVATDHAEILDRLIAGARTDLKDDARTTAADIA 303
RESULT 13
ID KBF2_HUMAN
AC Q00653;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR NF-KAPPA-B p100 SUBUNIT (H2TF1) (ONCOGENE LYT-10)
DE (LYT10) [CONTAINS: NUCLEAR FACTOR NF-KAPPA-B P52 SUBUNIT].
OS Homo sapiens (Human).
GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=91343004; PubMed=1876189;
RA Schmid R.M., Perkins N.D., Duckett C.S., Andrews P.C., Nabel G.J.;
RT "Cloning of an NF-kappa B subunit which stimulates HIV transcription
RT in synergy with p65."
RL Nature 352:733-736(1991).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=92123193; PubMed=1531086;
RA Bours V., Burd P.R., Brown K., Villalobos J., Park S., Ryseck R.P.,
RA Bravo R., Kelly K., Siebenlist U.;
RT "A novel mitogen-inducible gene product related to p50/p105-NF-kappa
RT B participates in transactivation through a kappa B site."
RL Mol. Cell. Biol. 12:685-695(1992).
[3]
SEQUENCE FROM N.A.
RX MEDLINE=94309910; PubMed=8036016;
RA Thakur S., Lin H.C., Tseng W.T., Kumar S., Bravo R., Foss F.,
RA Gelinas C., Rabson A.B.;
RT "Rearrangement and altered expression of the NFKB-2 gene in human
RT cutaneous T-lymphoma cells."
RL Oncogene 9:2335-2344(1994).
[4]
SEQUENCE OF 1-220 FROM N.A.
RX MEDLINE=95059001; PubMed=7969113;
RA Liptay S., Schmid R.M., Nabel E.G., Nabel G.J.;
RT "Transcriptional regulation of NF-kappa B2: evidence for kappa B-
RT mediated positive and negative autoregulation."
RL Mol. Cell. Biol. 14:7695-7703(1994).
[5]
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 37-327.
RX MEDLINE=98046037; PubMed=9384586;
RA Cramer P., Larson C.J., Verdine G.L., Mueller C.W.;
RT "Structure of the human NF-kappaB p52 homodimer-DNA complex at 2.1-A
RT resolution."
RL EMBO J. 16:7078-7090(1997).
-1- FUNCTION: P100 IS THE PRECURSOR OF THE P52 SUBUNIT OF THE NUCLEAR
FACTOR NF-KAPPA-B, WHICH BINDS TO THE KAPPA-B CONSENSUS SEQUENCE
5'-GGGNNYYCC-3', LOCATED IN THE ENHANCER REGION OF GENES INVOLVED
IN IMMUNE RESPONSE AND ACUTE PHASE REACTIONS. THE PRECURSOR
PROTEIN ITSELF DOES NOT BIND TO DNA.
-1- SUBUNIT: ACTIVE NF-KAPPA-B IS A HETERODIMER OF AN ABOUT 52 KDA
DNA-BINDING SUBUNIT AND THE WEAK DNA-BINDING SUBUNIT P65. TWO
HETERODIMERS MIGHT FORM A LABILE TETRAMER.
-1- SUBCELLULAR LOCATION: NUCLEAR, BUT ALSO FOUND IN THE CYTOPLASM
IN AN INACTIVE FORM COMPLEXED TO AN INHIBITOR (I-KAPPA-B).
-1- ALTERNATIVE PRODUCTS: NFKB2 CAN PRODUCE BOTH A 100 KDA AND A 49
KDA PRODUCT FROM AN ALTERNATIVELY SPLICED TRANSCRIPT.
-1- DOMAIN: THE C-TERMINUS OF P100 MIGHT BE INVOLVED IN CYTOPLASMIC
RETENTION, INHIBITION OF DNA-BINDING BY P52 HOMODIMERS, AND/OR
TRANSCRIPTION ACTIVATION (BY SIMILARITY).
-1- DISEASE: IN CUTANEOUS T-CELL LEUKEMIA (CTCL) CELL LINE HUT 78, A
REARRANGED NFKB2 GENE IS FOUND WHICH ENCODES FOR A TRUNCATED 80
KDA PROTEIN (P80HT).
-1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
-1- SIMILARITY: CONTAINS 7 ANK REPEATS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X61498; CAA43715.1; -
EMBL; S76638; AAB21124.1; -
EMBL; U09609; AAA21462.1; -
EMBL; U20816; AAA68171.1; -
PIR; S17233; S17233.
```

PDB: 1A30; 27-MAY-98.
DR TRANSFAC: T01927; -
DR MM: 164012; -
DR InterPro: IPR000451; -
DR InterPro: IPR000488; -
DR InterPro: IPR002110; -
DR Pfam: PF00554; RHD: 1.
DR Pfam: PF00023; ank: 6.
DR Pfam: PF00531; death: 1.
DR PRINTS: PRO0057; NFKB1NSCPFT.
DR PROSITE: PS0088; ANK_REPEAT; 5.
DR PROSITE: PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS01204; REL_1; 1.
DR PROSITE: PS0254; REL_2; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW Repeat; ANK repeat; Phosphorylation; Alternative splicing;
KW Disease mutation; Proto-oncogene; 3D-structure.
FT DOMAIN 38 343
FT REPEAT 486 518 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 558 590 ANK 3.
FT REPEAT 598 627 ANK 4.
FT REPEAT 632 662 ANK 5.
FT REPEAT 666 695 ANK 6.
FT REPEAT 728 757 ANK 7.
FT DOMAIN 337 341
FT DOMAIN 350 400
FT VARIANT 666 666
FT VARIANT 667 898
FT VARIANT 144 144
FT CONFLICT 213 213
FT CONFLICT 396 396
FT CONFLICT 433 434
FT CONFLICT 455 468
FT CONFLICT 740 740
FT CONFLICT 858 858
FT CONFLICT 874 874
FT CONFLICT 889 889
FT CONFLICT 898 898
SQ SEQUENCE 898 AA; 96752 MW; DFD153EB9E23A42B CRC64;
Query Match 20.5%; Score 224.5; DB 1; Length 898;
Best Local Similarity 33.9%; Pred. No. 2e-12;
Matches 65; Conservative 29; Mismatches 85; Indels 13; Gaps 5;
Qy 16 DKSLATRTDQSRTHLWACSGAHTIEVFLQLQGVFVNDKDDAGWSPLHIA--ASAGR 73
Db 515 DLGVVNLTHLHQTPPLHLAVITGTSVSVFLRVGADPALLDRHGDSAMHLALRAGAGAP 574
Qy 74 EIVKALLKGAQVNAVQ-----NCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
Db 575 ELLRALLQSGAP--AVPQLLHMPDFEGLYPVHLAVARSPECIDLLYDSCAEVEATERQ 632
Qy 127 A-TAMHRAAAGNLKMI-HILLYYKASTNIQDTGNTPLHLACDEERVEAKLLVSGAS 184
Db 633 GRTALHLATEMEELGLVTHVTKLRANVARTFAGNTPLHLAAGLYPTLRLLLKAGAD 692
Qy 185 IYIENKEETPL 196
Db 693 IHAENEPLCLPL 704
RESULT 14
ID PA26_HUMAN STANDARD; PRT; 806 AA.
AC O60733; O75645; Q9Y671; Q9U010; Q9UG29;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
101-OCT-2000 (Rel. 40, Last annotation update)
85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-PLA2).
PLA2G6 OR IPLA2.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID:9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS LH-IPLA; ANKYRIN-IPLA2-1 AND -2).
TISSUE=B-cell, and Testis;
MEDLINE=98079046; PubMed=9417066;
Larsson P.K.A., Claesson H.-E., Kennedy B.P.;
"Multiple splice variants of the human calcium-independent phospholipase A2 and their effect on enzyme activity.";
J. Biol. Chem. 273:207-214(1998).
[2]
SEQUENCE FROM N.A. (ISOFORMS LH-IPLA AND SH-IPLA).
TISSUE=Pancratic islets;
MEDLINE=99194813; PubMed=10092647;
Ma Z., Wang X., Nowatzke W., Ramanadham S., Turk J.;
"Human pancreatic islets express mRNA species encoding two distinct catalytically active isoforms of group VI phospholipase A2 (IPLA2) that arise from an exon-skipping mechanism of alternative splicing of the transcript from the IPLA2 gene on chromosome 22q13.1.";
J. Biol. Chem. 274:9607-9616(1999).
[3]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
MEDLINE=99269033; PubMed=10336645;
Larsson Forsell P.K.A., Kennedy B.P., Claesson H.-E.;
"The human calcium-independent phospholipase A2 gene. Multiple enzymes with distinct properties from a single gene.";
Eur. J. Biochem. 262:575-585(1999).
[4]
SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
Graham D.;
Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
TISSUE=Testis;
Anson W., Winkler U., Mewes H.-W., Gassenhuber J., Wiemann S.;
Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM PHOSPHOLIPIDS. IT HAS BEEN IMPLICATED IN NORMAL PHOSPHOLIPID REMODELLING, NITRIC OXIDE-INDUCED OR VASOPRESSIN-INDUCED ARACHIDONIC ACID RELEASE AND IN LEUKOTRIENE AND PROSTAGLANDIN PRODUCTION. MAY PARTICIPATE IN FAS MEDIATED APOPTOSIS AND IN REGULATING TRANSMEMBRANE ION FLUX IN GLUCOSE-STIMULATED B-CELLS.
-!- FUNCTION: ISOFORMS ANKYRIN-IPLA2-1 AND -2, WHICH LACK THE CATALYTIC DOMAIN, ARE PROBABLY INVOLVED IN THE NEGATIVE REGULATION OF IPLA2 ACTIVITY.
-!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-PHOSPHOCHOLINE + A FATTY ACID ANION.
-!- SUBUNIT: FORMS LARGE OLIGOMERIC 270-350 KDA STRUCTURES.
-!- SUBCELLULAR LOCATION: THE LONG ISOFORM WAS FOUND TO BE MEMBRANE BOUND. THE SHORT ISOFORM IS CYTOPLASMIC.
-!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; LH-IPLA2 (SHOWN HERE), SH-IPLA2, ANKYRIN-IPLA2-1 AND ANKYRIN-IPLA2-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- TISSUE SPECIFICITY: FOUR DIFFERENT TRANSCRIPTS WERE FOUND TO BE EXPRESSED IN A DISTINCT TISSUE DISTRIBUTION.
-!- SIMILARITY: CONTAINS 7 ANK REPEATS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF064594; AAC97486.1; -
EMBL; AF102988; AAD41722.1; -

Query Match	20.4%;	Score 222.5;	DB 1;	Length 806;
Best Local Similarity	27.9%;	Pred. No. 2.7e-12;		
Matches 62;	Conservative 40;	Mismatches 81;	Indels 39;	Gaps 5;

Search completed: August 13, 2001, 07:44:52
Job time: 29 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:47:24 ; Search time 32.54 Seconds
(without alignments)
866.041 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AYSGLLEELKESILADKSLA.....TPLQVAKGGLILKRMVEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1093	100.0	226	4	O75832
2	1037	94.9	231	11	O922X2
3	1035	94.7	231	11	O922X3
4	511.5	46.8	118	4	O85533
5	338	30.9	247	10	O92Q79
6	312.5	28.6	1943	11	O61307
7	310	28.4	237	3	O9P3P6
8	310	28.4	1088	4	O13484
9	310	28.4	4377	4	O12955
10	306	28.0	2622	11	O70511
11	304.5	27.9	636	5	O9T267
12	301	27.5	1762	11	O88521
13	297	27.2	1159	5	O9NCP8
14	297	27.2	2443	5	O9VSA2
15	290.5	26.6	1136	6	O9N180
16	288.5	26.4	1786	5	O17344
17	288.5	26.4	1809	5	O17487
18	288.5	26.4	1815	5	O17488
19	288.5	26.4	1867	5	O17486

20	288.5	26.4	2039	5	Q17489
21	288.5	26.4	6994	5	Q17343
22	287.5	26.3	1848	11	O61302
23	286.5	26.2	1719	4	Q13768
24	286.5	26.2	1856	4	Q99407
25	286	26.2	843	11	P97582
26	283.5	25.9	1059	4	O15084
27	283.5	25.9	1166	4	O9H2K2
28	283.5	25.9	1265	4	O9HASA
29	282	25.8	1098	11	O61304
30	281.5	25.8	1327	4	O95271
31	281	25.7	1181	5	O9XZ37
32	281	25.7	1181	5	O9VBP3
33	276	25.3	1549	5	O24241
34	276	25.3	1549	5	O9V4B1
35	269.5	24.7	815	6	O9TV77
36	269.5	24.7	1030	4	O14974
37	268.5	24.6	658	11	O62937
38	268.5	24.6	976	11	Q10728
39	267	24.4	963	13	Q90624
40	267	24.4	1004	13	Q90623
41	266	24.3	768	4	O9ULJ7
42	260.5	23.8	1062	11	O89019
43	260	23.8	1430	11	O9JJP7
44	259.5	23.7	1395	5	O9XZC0
45	258.5	23.7	2119	5	O9VAU5

ALIGNMENTS

RESULT 1

O75832 PRELIMINARY; PRT; 226 AA.

AC O75832; 01-NOV-1998 (TREMREL. 08, Created)

DT 01-NOV-1998 (TREMREL. 08, Last sequence update)

DT 01-MAR-2001 (TREMREL. 16, Last annotation update)

DE 26S PROTEASOME SUBUNIT P28,

GN DJ889N15.2 OR HUMAN GANKYRIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98382579; PubMed=9714768;

RA Hori T., Kato S., Saeki M., DeMartino G.N., Slaughter C.A.,

RA Takeuchi J., Toh-E A., Tanaka K.;

RT "cDNA cloning and functional analysis of p28 (Nas6p) and p40.5

RT (Nas7p), two novel regulatory subunits of the 26S proteasome.";

RL Gene 216:113-122(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Grafham D.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RT Higashitsuji H., Fujita J.;

RT "Enhanced expression of a novel tumour marker in the human

RT hepatomas.";

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB009619; BAA33215.1; -

DR EMBL; AL031177; CAA20117.1; -

DR EMBL; D83197; BAA34594.1; -

DR HSSP; P42773; 1IHB.

DR InterPro; IPR002110; -

DR Pfam; PF00023; ank; 5.

DR PROSITE; PS50086; ANK_REPEAT; 5.

DR PROSITE; PS50297; ANK_REPEAT; 1.

DR SMART; SM00248; ANK; 1.

KW Proteasome.

```
SQ SEQUENCE 226 AA; 24428 MW; 57158E33146EC7C8 CRC64;

Query Match 100.0%; Score 1093; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 6.7e-91;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRMVEG 213
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRMVEG 226

RESULT 2
ID Q922X2 PRELIMINARY; PRT; 231 AA.
AC Q922X2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GANKYRIN.
GN PSMD10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of mouse gankyrin containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022022; BAA36959.1; -.
DR HSSP; P42773; 1IHB.
DR MGP; MGI-1858898; Psmd10.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 25115 MW; 5A3DB7027B0594E6 CRC64;

Query Match 94.9%; Score 1037; DB 11; Length 231;
Best Local Similarity 93.9%; Pred. No. 7.7e-86;
Matches 199; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLGKGAHNSVQNGCTPLHYAASKNRHEISVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRMVE 212
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRLAE 225
```

```
RESULT 3
Q922X3 PRELIMINARY; PRT; 231 AA.
ID Q922X3;
AC Q922X3;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GANKYRIN HOMOLOGUE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022014; BAA36954.1; -.
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;

Query Match 94.7%; Score 1035; DB 11; Length 231;
Best Local Similarity 94.3%; Pred. No. 1.2e-85;
Matches 200; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLQGVPNVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLIKGAQVNAVQNGCTALHYAASKNRHEIAVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKHYDATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVT 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRMVE 212
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRLAE 225

RESULT 4
Q95533 PRELIMINARY; PRT; 118 AA.
ID Q95533;
AC Q95533;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE DJ889N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
DE (PUTATIVE PARTIAL ISOFORM 2) (FRAGMENT).
GN DJ889N15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031177; CAA20118.1; -.
DR HSSP; Q13625; IYCS.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 2.
```

Qy 179 VSGASTIYENKEEKPLOVAKGLGLIL- ----KRWEG 213
:
Db 209 IIRGADVDEBKGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 6
Q61307 PRELIMINARY; PRT; 1943 AA.
ID Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; O08866; O08867;
AC 01-JUN-1998 (TEMBLrel. 06, Created)
DT DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE ANKYRIN 3 (ANKYRIN G) (EPITHELIAL ANKYRIN) (ANKYRIN-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

[1]
RN SEQUENCE FROM N.A. (ISOFORMS 1-6).
RP STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=95340633; PubMed=715634;
RA Peters L.C., John K.M., Lu F.M., Eichler E.M., Higgins A., Yialamas M.,
RT Turtso L.C., Otsuka A.J., Lux S.E.;
RT "Ank3 (epithelial ankrynin), a widely distributed new member of the
RT ankrynin gene family and the major ankrynin in kidney, is expressed in
RT alternatively spliced forms, including forms that lack the repeat
RT domain.";
RL J. Cell Biol. 130:313-330(1995).
RN [2]

RN SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RP STRAIN=C3H/HEJ; TISSUE=BONE MARROW;
RX MEDLINE=97213781; PubMed=9060470;
RA Hoock T.C., Peters L.L., Lux S.E.;
RT "Isoforms of ankrynin-3 that lack the NH2-terminal repeats associate
RT with mouse macrophage lysosomes.";
RL J. Cell Biol. 136:1059-1070(1997).

CC -1- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
SKELLETAL ELEMENTS; THEY BIND TO THE ERTHROCYTE MEMBRANE PROTEIN GP85,
BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCTE MEMBRANE PROTEIN GP85,
AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
DESMIN. ERTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
THE CYTOPLASMIC DOMAIN OF THE ERTHROCYTES ANION EXCHANGE PROTEIN.
CC THEY REPEAT MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
ARE ASSOCIATED WITH THE LYXOSOMAL MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,
4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES, IN
LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC -1- SIMILARITY: CONTAINS 24 ANK REPEATS.

DR ENBL; L40631; AAB01603.1; -
DR ENBL; L40632; AAB01605.1; -
DR ENBL; U89275; AAB58381.1; -
DR ENBL; L40631; AAB01602.1; -
DR ENBL; L40632; AAB01604.1; -
DR ENBL; L40632; AAB01606.1; -
DR ENBL; L40632; AAB01607.1; -
DR ENBL; U89274; AAB58380.1; -
DR HSP; P55273; 1BD8.
DR MGD; MG1:88026; Ank3.
DR InterPro; IPR000488; -
DR InterPro; IPR000906; -
DR InterPro; IPR021110; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00005; DEATH; 1.

KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Multigene family.
FT DOMAIN 1 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT DOMAIN 23 810 24 X ANK MOTIF REPEATS.
FT DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.
FT DOMAIN 1438 1943 REGULATORY DOMAIN (REGULATES BINDING OF
ANKYRIN TO SPECTRIN AND THE BAND 3
PROTEIN).
FT REPEAT 23 55 ANK MOTIF 1.
FT REPEAT 56 88 ANK MOTIF 2.
FT REPEAT 89 121 ANK MOTIF 3.
FT REPEAT 122 154 ANK MOTIF 4.
FT REPEAT 155 183 ANK MOTIF 5.
FT REPEAT 184 216 ANK MOTIF 6.
FT REPEAT 217 249 ANK MOTIF 7.
FT REPEAT 250 282 ANK MOTIF 8.
FT REPEAT 315 348 ANK MOTIF 9.
FT REPEAT 349 381 ANK MOTIF 10.
FT REPEAT 382 414 ANK MOTIF 11.
FT REPEAT 415 447 ANK MOTIF 12.
FT REPEAT 448 480 ANK MOTIF 13.
FT REPEAT 481 513 ANK MOTIF 14.
FT REPEAT 514 546 ANK MOTIF 15.
FT REPEAT 547 579 ANK MOTIF 16.
FT REPEAT 580 612 ANK MOTIF 17.
FT REPEAT 613 645 ANK MOTIF 18.
FT REPEAT 646 678 ANK MOTIF 19.
FT REPEAT 679 711 ANK MOTIF 20.
FT REPEAT 712 744 ANK MOTIF 21.
FT REPEAT 745 777 ANK MOTIF 22.
FT REPEAT 778 810 ANK MOTIF 23.
FT VARSPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 834 855 VRKASPEKLSGDEYISDGEEG -> GDKCTWFKPKVQEV
LVK (IN ISOFORM 5).
FT VARSPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND
ISOFORM 3).
FT VARSPLIC 852 852 G -> GDKCTWFKPKVQEVLVKS (IN ISOFORM 4
AND ISOFORM 6).
FT VARSPLIC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND
ISOFORM 7).
SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

Query Match 28.6%; Score 312.5; DB 11; Length 1943;
Best Local Similarity 35.5%; Pred. No. 2.9e-19;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSAGTEIVEFLLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82
DB 517 TPLHIAAREGHEDVAFLLDHGASLSITTKGFTPLHVAAYKGLVAVSLLQKSASPD 576
QY 83 -----GAQVNAVNGCTGPLHYAASKNRHEIAVMLEGG 115
DB 577 AGKSGLTPLHVAAYHNDGKVALLLDQASPHAAAKNGYTPHIAAKKNQMDIATSLLEY 636
QY 116 GANPAKDHYEATAMHRAAAGNLKMIHILLYKASTNIQDTEGNTPLHLACDEERVEEA 175
DB 637 GADANAVTQGTASVHAAQEGHVDMSLLSRNANVNLNKSGLTPLHLAAQEDRVNVA 696
QY 176 KLVVSGASIVYENKEETPLQV 198
DB 697 EVLVNQAHAVDQATKMGYTPLVH 719

RESULT 7
ID Q9P3P6 PRELIMINARY; PRT; 237 AA.
AC Q9P3P6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RELATED TO 26S PROTEASOME SUBUNIT P28.
GN B7F18.30.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL398981; CAB97304.1;
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 6.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome. 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;
SQ SEQUENCE 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;

Query Match 28.4%; Score 310; DB 3; Length 237;
Best Local Similarity 34.5%; Pred. No. 2.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

QY 1 AYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGTEIVEFLL-QLGVVNDKDDA 59
DB 12 ARDGKASIVESLLNANPKLAQRKDDGRLPIHWACSYNRKVEYVLLVNGKDFDPVEDDM 71
QY 60 GWSPLHIAASA-GRDEIVKALLGGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 118
DB 72 GMTFFMISAVSDSDAIILLRSRGADINQTNHQSOTALHFTASKNNIDILARKLLSPDMK 131
QY 119 PD-----AKDHVEATAMHRAAAGNLKMIHILLYKASTNIQDTEGNTPLHLACDEERVE 173
DB 132 PKPASVRVNDKRGQYPLHRAAAGISVPMINLLQHKSPINASDNAGYTPHHLHVAEGHGH 191
QY 174 EAKLVVSGASIVYENKEETPLQV 198
DB 192 AAVALLKAGAEKTKDKMDGYLALDLAPDKDVRFRFIEKEAKEGIEL 237

RESULT 8
ID Q13484 PRELIMINARY; PRT; 1088 AA.
AC Q13484;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ANKYRIN G119.
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=96234670; Pubmed=8666667;
RA Kashgarian M., Morrow J.S., Ardito T., Stabach P.R., Mann A.S.,
RA Devarajan P.;
RT "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
and muscle that binds beta I sigma spectrin and associates with the
Golgi apparatus";
RL J. Cell Biol. 133:819-830(1996).
DR EMBL; U43965; AAB08437.1;
DR HSP; P55273; IBD8.
DR InterPro; IPR000906;

```

DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 12.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE; PS5088; ANK_REPEAT; 12.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00218; ZU5; 1.
DR SMART; SM00218; ZU5; 1.
DR SQ SEQUENCE 1088 AA; 119427 MW; 769C88D40A78DE86 CRC64;

```

```
Query Match      28.4%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 2.2e-19;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;
```

Qy	7	BELKESLIADKSLATRDQDSRTALHWASAGHTEIVEFLQLGVPNVDKDDAGWSPLHI	66
Db	166	EDVAAFLDDHGASLSITTKKGFTPLHVAARYGLEYVANLLQKASPDAAKGSGLTPLHV	225
Qy	67	AASAGRDEIVKALLGGCAOVNAVNOGCTPLHVAASKNRHEIATVMLLEGANPDADKHYE	126
Db	226	AAHYDNQVALLDDOGASPHAAKNGYTPLHIAAKNKQMDIATTLLEYGADANAVTROG	285
Qy	127	ATAMHRAAAAGNLKMTIHILYYKASTNIODTEGNTPLHLACDDEERVEEAKLIVSQGASY	186
Db	286	IASVHLAAQEGSHVDYSLLLGRNAVYNLSNKSGLTPLHLAAQEDRYNVAEVLVNOGAHYD	345
Qy	187	IENKEETPLOV	198
Db	346	AOTKMGYTPLHV	357

RESULT	9	
Q12955		
ID	Q12955	PRELIMINARY;
AC	Q12955;	PRT; 4377 AA.
DT	01-NOV-1996	(TREMBlrel. 01, Created)
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)
DT	01-MAR-2001	(TREMBlrel. 16, Last annotation update)
DE	ANKYRIN G.	
GN	ANK-3.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	NCBI TaxID=9606.	

OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN STEM;
RX	MEDLINE=95138209; PubMed=7836459;
RA	Kordeli E., Lambert S., Bennett V.;
RT	"Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT	at the axonal initial segment and node of Ranvier.";
RJ	J. Biol. Chem. 270:2352-2359(1995).

14]	SEQUENCE FROM N.A.	
RP	TISSUE-BRAIN STEM;	
RC	Carpenter S.S.;	
RA	Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.	
RL	EMBL: U13166; AAA64834.1; -	
DR	DB	
DR	HSSP: P55273; 1BD8	
DR	InterPro: IPR000488; -	
DR	InterPro: IPR000906; -	
DR	InterPro: IPR002110; -	
DR	Pfam: PF00023; ank; 23	
DR	Pfam: PF00531; death; 1	
DR	Pfam: PF00791; zUS; 1	
DR	PROSITE: PS50088; ANK_REPEAT; 21	
DR	PROSITE: PS50297; ANK_REP_REGION; 1	
DR	PROSITE: PS50017; DEATH_DOMAIN; 1	
DR	SMART: SM00005; DEATH; 1	
SQ	SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;	

Query Match 28.4%; Score 310; DB 4; Length 4377;
Best Local Similarity 37.0%; Pred. No. 1.4e-18;

	Matches	71:	Conservative	36:	Mismatches	85:	Indels	0:	Gaps	0:
Qy	7	BELKESILADRSKLATRTDQDSRTALHWACSSAGHTEIVFLLQLGVPVNDKDDAGWSPLHI	66	: :	:	: :				
Db	545	EDVAAFLLDHGCASLSTITTKGFTHLVAAKYGVKEVANLLLOKSASPDAACKSGLTPLHV	604	: :	:	: :				
Qy	67	AASAGRDEIVKALLKGQAQNVNAVNWNGCTPLHYAASKNRHEIAVWLLLEGGANPDRAKHYE	126	: :	:	: :				
Db	605	AAHYDNOKVALLLDQGASHPAHAANKGYTPPHIAAKNQMDIATTLLEYGDADANAVTROG	664	: :	:	: :				
Qy	127	ATAMHRAAAKNLKMITHLYIKASTNIDTFEGNTPTPLACLADEERVEEAKLLVSOGASIY	186	: :	:	: :				
Db	665	IASYHLAAQEGHDVMSLLLRGNANVLNSKSGLTPLHLAAQEDRVNVAEVYNQGAHVD	724	: :	:	: :				
Qy	187	TENKEEKTPLQV	198	:						
Db	725	AQTKMGYTPLHV	736	:						

```

RESULT 10
O70511
ID O70511 PRELIMINARY; PRT; 2622 AA.
AC O70511;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1999 (TREMBlrel. 11, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE 270 KDA ANKYRIN G ISOFORM (ANKYRING) (FRAGMENT).
GN ANK3.

```

OS *Rattus norvegicus* (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98417633; PubMed=9744885;
 RX Zhang X., Bennett V.;
 RA "Restriction of 480/270-kD ankyrin G to axon proximal segments
 RT requires multiple ankyrin G-specific domains";
 RL J. Cell Biol. 142:1571-1581(1998).

```

[3] SEQUENCE FROM N.A.
RP Carpenter S.S., Zhang X.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL
RN
[3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RL J. Cell Sci., 0:0-0(1998).
DR EMBL; AF102552; AAC78143.1; -.
DR EMBL; AF065150; AAC18853.1; -.
DR HSSP; P55273; 1BD8.
DR InterPro; IPR000486; -.
DR InterPro; IPR000906; -.
DR InterPro; IPR02110; -.
DR Pfam; PF00203; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

```

```
Query Match      28.0%; Score 306; DB 11; Length 2622;
Best Local Similarity 36.5%; Pred. No. 1.7e-18;
Matches 70; Conservative 36; Mismatches 86; Indels 0; Gaps 0;
```

Db 553 EDVAEFLDHGASLSITTKKGTPLHVAAYKGLVLEASLLLOKSASPDAAAGSGLTPLHV 612
QY 67 AASAGRDEIVKALIGKAQVNAVONQCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
Db 613 AAHYDNOKVALLLDQOGASPHAAKNGYTPHIAAKKNQMDIATSLLEYGADANPYTROG 672
QY 127 ATAMHRAAAKGNLMKIHILLYYKASTNIQDTEGNTPLHACDERVEEAKLLYSOGASIV 186
Db 673 IASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 732
QY 187 IENKEEKTPLQV 198
Db 733 AQTKMGYTPLVH 744
RESULT 11
Q9TZ67 PRELIMINARY; PRT: 636 AA.
AC Q9TZ67;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE F40G9.1 PROTEIN.
GN F40G9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.";
RN Nature 368:32-38(1994).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RA Graves T., Sutterer C., Ozersky P.;
RT "The sequence of C. elegans cosmid F40G9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099919; AAC68798.1; -;
DR HSP; Q00421; IAWC.
DR InterPro: IPR002110; -;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 636 AA; 73229 MW; AB55162AF5D5B1C0 CRC64;
Query Match 27.9%; Score 304.5; DB 5; Length 636;
Best Local Similarity 34.6%; Pred. No. 3.4e-19;
Matches 85; Conservative 39; Mismatches 75; Indels 47; Gaps 7;
QY 1 AYSGLKEELKESILADKSLATRTDQDSRTAL-----HWACSAGH 39
Db 379 AVGGSILPQLQFAINDPEMAHKTD-DVRVGMWNLTEIENFFNRKICDFRKNYSKNTKN 437

QY 40 TEIVFELLO-----LGVPVNDK-----DAGWSPHLHIAASAGRDEIVK 77
Db 438 PFSEFFETIDFEHENSONALKKFSLKIDIFDLNCLILPLGWTPLMIASSAGRDVVR 497
QY 78 ALLG-KGAQVNAVONQCTPLHYAASKNRHEIAVMLEGGAN-PDAKDHYEATAMHRAAA 135
Db 498 YLLTLPDQVVKUTNSNKQTSLVYACSKNHEIVKLLIEADPNILNLPDFEGATALHRAAS 557
QY 136 KGNLMKIHILLYY-KASTNIQDTEGNTPLHACDERVEEAKLLYSOGASIVYENKEEKT 194
Db 558 RGNQDVIIVRALVSTGKSLRDQEGEGTALHLACDENRGDVAILLVNRGADMKNLKEKOT 617
QY 195 PLQVAK 200
Db 618 PLEMLK 623
RESULT 12
O88521 PRELIMINARY; PRT: 1762 AA.
ID O88521;
AC O88521;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE 190 KDA ANKYRIN ISOFORM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kollu A.H., Devatajan P.;
RT Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, K-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -;
DR HSP; P55273; 1B18.
DR InterPro: IPR000169; -;
DR InterPro: IPR000488; -;
DR InterPro: IPR000906; -;
DR InterPro: IPR002110; -;
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; UNKNOWN_1.
DR SMART; SM00005; DEATH; 1.
SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;
Query Match 27.5%; Score 301; DB 11; Length 1762;
Best Local Similarity 35.9%; Pred. No. 2.8e-18;
Matches 69; Conservative 36; Mismatches 87; Indels 0; Gaps 0;
QY 7 EELKESILADKSLATRTDQDSRTALHWACSAGHTEIVFELIQLGVVPVNDKDGWSPHLHI 66
Db 539 EDVAEFLDHGAFSLITTKKGTPLHVAAYKGLVLEASLLLOKSASPDAAAGSGLTPLHV 598
QY 67 AASAGRDEIVKALIGKAQVNAVONQCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
Db 599 AAHYDNOKVALLLDQOGASPHAAKNGYTPHIAAKKNQMDIATSLLEYGADANAVTROG 658
QY 127 ATAMHRAAAKGNLMKIHILLYYKASTNIQDTEGNTPLHACDERVEEAKLLYSOGASIV 186
Db 659 IASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 718
QY 187 IENKEEKTPLQV 198

```

Db 719 AOTRMGYTPLHV 730
RESULT 13
Q9NCP8 PRELIMINARY; PRT; 1159 AA.
AC Q9NCP8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ANKYRIN 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The li-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL; AF190635; AAF73309.1; -
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000906; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

Query Match 27.2%; Score 297; DB 5; Length 1159;
Best Local Similarity 29.7%; Pred. No. 3.6e-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 12 SLADK--SLATRTDQDSRTALHWACSAAGTEIVEFLLQLGVVNDKDDAGWSP----- 63
Db 255 SLLLEKGGNIEAKT-RDGLTPLHCAARSGHQVVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 64 -----LHIAASAGRDEIVKALLGKGAQVNAVNGCTP 96
Db 314 GEHVDAARILLYHRAPVDEVTVDTALHVAACHGVHVAKLLLDNRNADANARALNGFTP 373
Qy 97 LHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQD 156
Db 374 LHIAACKNRKLVVLELLRHGASISATSGTLPLHVAAFMGCMNIVLYLQHDASPDVPT 433
Qy 157 TEGNTPPLHACDEEREEAKLVSGASIIYENKEKTPLOVAK--GGIGLILKRWVSG 213
Db 434 VRGETPLHARANGTDIIRILLRNGAQVDARARQQTPPLHTASRLGNVDIVMLLQHG 492

RESULT 14
Q9VSA2 PRELIMINARY; PRT; 2443 AA.
AC Q9VSA2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ESTS:149B105 PROTEIN.
GN ANK2 OR ESTS:149B105 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.N., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector A., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003558; AAF50525.1; -.
DR HSSP; P42773; 1IHB.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 2443 AA; 269538 MW; 7561DC5CF56812DC CRC64;

```

```

Query Match 27.2%; Score 297; DB 5; Length 2443;
Best Local Similarity 29.7%; Pred. No. 9.8e-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 12 SLADK--SLATRTDQDSRTALHWACSAAGTEIVEFLLQLGVVNDKDDAGWSP----- 63
Db 255 SLLLEKGGNIEAKT-RDGLTPLHCAARSGHQVVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 64 -----LHIAASAGRDEIVKALLGKGAQVNAVNGCTP 96
Db 314 GEHVDAARILLYHRAPVDEVTVDTALHVAACHGVHVAKLLLDNRNADANARALNGFTP 373
Qy 97 LHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQD 156
Db 374 LHIAACKNRKLVVLELLRHGASISATSGTLPLHVAAFMGCMNIVLYLQHDASPDVPT 433

```

Qy 157 TEGNTPHLACDEERVEEAKLLVSGASIYIENKEETPLQVAK--GGGLGLILKRMVEG 213
Db 434 VRGETPLHLAARANQTDIIKILLRNGAQVDARAREQQTPLHTASRLGNVDIVMLLQHG 492

RESULT 15

Q9N180
ID Q9N180 PRELIMINARY; PRT; 1136 AA.
AC Q9N180;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ANKYRIN 1 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRED JAPANESE BLACK; TISSUE-BONE MARROW;
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT "Cloning of bovine erythrocyte ankyrin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF222766; AAF61702.1; -
DR InterPro; IPR000906; -
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
FT NON_TER 1136 1136
SQ SEQUENCE 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;

Query Match 26.6%; Score 290.5; DB 6; Length 1136;
Best Local Similarity 37.5%; Pred. No. 1.4e-17;
Matches 72; Conservative 30; Mismatches 81; Indels 9; Gaps 2;
Qy 13 ILADKSLATRTDQDSR-----TALHWACSGAGHTEIVEFLQLGVVPVNDKDDAGWSPLHIA 67
Db 355 VLLDKG----AKPNSRALNGFTPLHIAKKNHIRVMELLLKMGASIDAVTESGLTPLHVA 410
Qy 68 ASAGRDEIVKALLCGQAVNQNGCTPLHYAASKNRHETAVMLLEGGANPDQAKDHVEA 127
Db 411 SFMGHPPIVKSLLQREASPNVSNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQ 470
Qy 128 TAMHRAAKGNLKMIIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSGASIYI 187
Db 471 TPLHCAARIGHTNNVKULLENNANPNLATTAGHTPLHIAAREGHVETALALLEKEASQTC 530
Qy 188 ENKEETPLQVA 199
Db 531 MTKRGFTPLHVA 542

Search completed: August 13, 2001, 07:47:25
Job time: 181 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 13, 2001, 07:45:23 ; Search time 1109.56 Seconds
(without alignments)
10873.539 Million cell updates/sec

Title: US-09-509-775-1
Perfect score: 780
Sequence: 1 tggtagactctaacgctg.....gaatggtgaaggttaaca 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl: *
- 1: gb_bal:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_bal:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rod:*
 - 22: em_htg_hum1:*
 - 23: em_htg_hum2:*
 - 24: em_htg_hum3:*
 - 25: em_htg_hum4:*
 - 26: em_htg_hum5:*
 - 27: em_htg_hum6:*
 - 28: em_htg_hum7:*
 - 29: em_htg_hum8:*
 - 30: em_htg_inv1:*
 - 31: em_htg_inv2:*
 - 32: em_htg_other:*
 - 33: em_htg_rod:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_v1:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_v11:*
- 59: gb_v12:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_rod:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	780	100.0	1544	91	D83197 Homo sapien
2	706	90.5	1468	10	E12804 cDNA encodi
3	706	90.5	1468	85	AB009619 Homo sapi
4	593	76.0	696	94	AB022014 Rattus no
5	584	74.9	696	94	AB022022 Mus muscu
6	562	72.1	180905	62	AC012149 Homo sapi
7	532	68.2	105334	92	HS914P20
8	488.6	62.6	175550	61	AC009863 Homo sapi

9	398.2	51.1	199882	90	AL354720	Human DNA
10	223.8	28.7	145483	74	AC072018	Homo sapi
11	211	27.1	123395	92	AL031177	Human DNA
12	211	27.1	178492	78	AL137065	Homo sapi
13	115.6	14.8	5069	6	DROANKY	L35601 Drosophila
14	115.6	14.8	10152	63	AC014567	AC014567 Drosophila
15	115.6	14.8	114334	5	AE003845	AE003845 Drosophila
16	109.4	14.0	4408	85	AB002377	Homo sapi
17	98.4	12.6	2529	95	RNU65916	U65916 Rattus norv
18	96	12.3	6278	93	HSRANK2	X56958 Homo sapien
19	96	12.3	12518	93	HSRANK440	Z26634 Homo sapien
20	93.4	12.0	5235	9	AR048831	AR048831 Sequence
21	93.4	12.0	5235	9	AR048832	AR048832 Sequence
22	93.4	12.0	5235	9	AR050426	AR050426 Sequence
23	93.4	12.0	5235	9	AR050427	AR050427 Sequence
24	93.4	12.0	5503	9	AR048829	AR048829 Sequence
25	93.4	12.0	5503	9	AR048830	AR048830 Sequence
26	93.4	12.0	5503	9	AR050424	AR050424 Sequence
27	93.4	12.0	5503	9	AR050425	AR050425 Sequence
28	89.8	11.5	4818	5	AF190635	AF190635 Drosophila
29	89.2	11.4	102552	65	AC019368	AC019368 Drosophila
30	89.2	11.4	273744	4	AE003558	AE003558 Drosophila
31	88.4	11.3	3454	97	HSU43965	U43965 Human ankyr
32	88.4	11.3	14770	97	HSU13616	U13616 Human ankyr
33	88	11.3	3520	94	MMANK1AC	X69065 M.musculus
34	88	11.3	6238	94	MUSANK1A	M84756 Mouse red c
35	88	11.3	8145	94	MMANK1AA	X69063 M.musculus
36	86.8	11.1	7084	94	MUSANK3B	L40632 Mus musculu
37	86.6	11.1	911	9	AR048821	AR048821 Sequence
38	86.6	11.1	911	9	AR048822	AR048822 Sequence
39	86.6	11.1	911	9	AR050416	AR050416 Sequence
40	86.6	11.1	911	9	AR050417	AR050417 Sequence
41	85.6	11.0	909	9	AR048823	AR048823 Sequence
42	85.6	11.0	909	9	AR048824	AR048824 Sequence
43	85.6	11.0	909	9	AR050418	AR050418 Sequence
44	85.6	11.0	909	9	AR050419	AR050419 Sequence
45	85.2	10.9	6192	97	HUMANK	M28880 Human eryth
ALIGNMENTS						
RESULT	1					
D83197						
LOCUS	D83197	1544 bp	mRNA	PRI	06-JAN-1999	
DEFINITION	Homo sapiens mRNA for ankyrin repeat protein, complete cds.					
ACCESSION	D83197					
VERSION	D83197.1	GI:3893154				
KEYWORDS	human gankyrin; ankyrin repeat protein.					
SOURCE	Homo sapiens female placenta cDNA to mRNA, clone:yf68b03.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1. (bases 1 to 1544)					
TITLE	Higashitsuji.H.					
JOURNAL	Direct Submission					
REFERENCE	Submitted (20-JAN-1996) to the DDBJ/EMBL/GenBank databases. Hiroaki Higashitsuji, Faculty of Medicine, Kyoto University, Clinical Molecular Biology, Shogoin Kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto 605, Japan (E-mail:higashitsuji@viru1.virus.kyoto-u.ac.jp, Tel:81-75-751-3753, Fax:81-75-751-3750)					
AUTHORS	2 (sites)					
TITLE	Higashitsuji.H. and Fujita,J.					
JOURNAL	Enhanced expression of a novel tumour marker in the human hepatomas Unpublished (1996)					
FEATURES	Location/Qualifiers					
source	1..1544					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="yf68b03"					
	/sex="female"					
	/tissue_type="placenta"					
gene	99..779					

CDS	/gene="human gankyrin"	99..779
	/gene="human gankyrin"	
	/codon_start=1	
	/product="ankyrin repeat protein"	
	/protein_id="BAA34594.1"	
	/db_xref="GI:3893155"	
	/translation="MEGCVSLNMCNLAISKLEELKESILADKSLATRTDQDSRTALHWACSAHTEIVFELLQIGVFNKDDAGNSPLHIAASAGRDEIVKALLGGAGQVNAV NONGCTPLHYAASKNRHEIAVMLEGGANPDAKDHYEATAMHAKNLKMHILLY YKASNIODTEGNTPLHLACDEERVEEAKLLVSQASITYIENKEKTPLOVAKGLGL ILKRVGEG"	
BASE COUNT	435 a	287 c 354 g 468 t
ORIGIN		
Query Match	100.0%;	Score 780; DB 91; Length 1544;
Best Local Similarity	100.0%;	Pred. No. 1.9e-226;
Matches 780; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	tgtgtgaagctctaacggtgttttgaactggcgtagccggagccggcgacgtgagcgggc 60
Db	3	TGCTGAAGCTCTAACGGCTGTTTGACTGGCTAGCGGCGCGGCGACGTGAGGCGGC 62
QY	61	gttgcctcgcgcagacaagtgtgtcgtgggacagcgaataaggaggggtgtgtcctaacta 120
Db	63	GTTGCTCGCGCGCAACAAGTAGTTGCTGGGACAGCAAAATGGAGGGGTGTGTCTAAACCTA 122
QY	121	atggtctgaacctggcctacacgggagagctggaagagttgaaggagagattctctggcc 180
Db	123	ATGGTCTGAACCTGCGCCCTACAGCGGGAAGCTGGAAGAGTTGAAGGAGAGTAGTTCTGGCC 182
QY	181	gataaatccctgctactagaactgaccggagcagacagacagactgcattgcactggggcatgc 240
Db	183	GATAAATCCCTGCTACTAGAACTGACACGAGCAGACAGAACTGCATTGCACCTGGCATGC 242
QY	241	tcagctggacatacacagaattgttgaatttttgcacacttgagtgagtgagtgagtgat 300
Db	243	TCAGCTGGACATACAGAAATTTTGAATTTTCTTCAACTTGGAGTGGCAGTGAATGAT 302
QY	301	aaagacgatgcaggttggtctctctctcattctgctggctctctgctggcggggatgagatt 360
Db	303	AAAGACGATGCAGGTGGTCTCTCTCTTCATTATTTGGCGCTTCTGCTGCCGGGATGAGATT 362
QY	361	gtaaaagccctctgggaaaagtgctcgaagtgaatgctgtcgaatcaatacaaatgctgtact 420
Db	363	GTAAGAGCCCTTCTGGAAAAGGTGCTCAAGTGAATGCTGTCAATCAAAATGGCTGTACT 422
QY	421	cccttacattatgcagcttcgaaaaacagcagcagcagcagcagcagcagcagcagcagc 480
Db	423	CCCTTACATTATGCAGCTCGAAAAACAGGCATGAGATCGCTGTCAFTGTTTACTTGGAGGC 482
QY	481	ggggctaataccagatgctaaggaccattatgaggtctacagcaatgcacccgggagcagcc 540
Db	483	GGGGCTAATCCAGATGCTAAGGACCATTTATGAGCTACAGCAATGCACCGGGCAGCAGCC 542
QY	541	aagggttaactggaagtgtatcatctctctctctctctctctctctctctctctctctctct 600
Db	543	AAGGGTAACCTGAAGATGATTTCATCTCTCTGTACTTACAAAGCATCCCAACATCTCAA 602
QY	601	gacactgaggttaacactcctctacacttagcctgtgtatgagagagagtggaagaagca 660
Db	603	GACACTGAGGGTACACTCTCTACACTTAGCCTGTGTATGAGGAGAGTGGGAAGCA 662
QY	661	aaactgctgtgtcccaagagcaagtattatcatgtgagataaagaagaagaagacaccc 720
Db	663	AAACTGCTGGTGTCCCAAGGAGCAAGTATTATACATTGAGAAATAAGAGAGAAAGACACC 722
QY	721	ctgcaagtggccaaagtggcctgggttttaatactcaagagaatgggtgaggttaaca 780
Db	723	CTGCAAGTGGCCAAAGGTGGCTGGGTTTAACTCAAGAGATGGTGGAGAGGTAAACA 782

YKASTNIQDTGNTPLHLACDERVEEAKLLVQSQASIIYIENKEKTPLOVAKGGLGL
ILKRWVEG"

polyA_signal 1453...1458

BASE COUNT 424 a 268 c 324 g 452 t

ORIGIN

Query Match 90.5%; Score 706; DB 85; Length 1468;
Best Local Similarity 100.0%; Pred. No. 6.7e-204;
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 aagtagtcttgagcagcgaataagaggggtgtgtgttaacacctaagtctgcaacct 134
DB 1 AAGTAGTCTGGGACAGCGAATGGAGGGGTGTGTCTAACCTAATGGTGCACCT 60
QY 135 ggcctacagcgggaagctggaagagtggaagagagattcttgcgcgataataatccctggc 194
DB 61 GGCCTACAGCGGGAAGCTGGAAGAGTGAAGGAGAGTATCTGGCGGATAAATCCCTGCG 120
QY 195 tactagaactgaccagcagcagcagaactcattgcactggcgactcagctggagacatac 254
DB 121 TACTAGAATGACAGGACAGCAGAACTGCATGTCACCTGGGATGCTCAGCTGGACATAC 180
QY 255 agaaattgtgaatttttctgcaacttgagtgccagtggaatgataaagacgatgcagg 314
DB 181 AGAAATGTTGAATTTTGTGCAACTTGGAGTGCAGTGAATGATAAAGACGATGCAGG 240
QY 315 ttggtctctctcatatgcggctctctgctggcggagtgagattgtaaaagccctct 374
DB 241 TTGGTCTCTCTTCAATATTCGGGCTTCTGCTGGCGGATGAGATTGTAAGGCCCTTCT 300
QY 375 ggaagaagtgctcaagtaagtctgctcaatcaaatggctgactcccttacattatgc 434
DB 301 GGAAGAAGTGCTCAAGTGAATGCTGTCAATCAAAATGCTGTACTCCCTTACATATGC 360
QY 435 agcttggaaaaacagcagcatgagatgcgtctgcttacttggaagcgggggctaataccaga 494
DB 361 AGCTTCGAAAAACAGCAGCATGAGATGCTGTCAATGTTACTTGAAGGGGGGCTAATCCAGA 420
QY 495 tctaaggaccattatgagctacacaaatgcacccggcagcagcccaaggttaacttga 554
DB 421 TGTAAGGACCATATGAGGCTACAGCAATGACACCGGGCAGCAGCCAAAGGTAACTTGAA 480
QY 555 gatgattcatctctctgctactacaagcatocacaacacatccaagacactgagggtaa 614
DB 481 GATGATTCAATCTTCTGTACTTACAAGCATCCACAACATCCAAAGACACTGAGGGTAA 540
QY 615 cactctctacattagcctgtgtatgagagagagtggaagaagcaaaactgctgggtgc 674
DB 541 CACTCTCTTACACTTACCTTACCTGTGATGAGGAGAGTGGAGAAAGCAAAACTGCTGGTGTG 600
QY 675 ccaaggagcaagattattacattagaataaagaaaaagacacccctcaagtggccaa 734
DB 601 CCAAGGAGCAAGTATTACATTGAGTAATAAGAAAGAAAGACACCCCTGCAAGTGCCCAA 660
QY 735 aggtggcctgggtttaaactcaagaagaatgggtggaaggttaaca 780
DB 661 AGGTGGCTTGGTTTAACTACTCAAGAGAAATGGTGAAGGTTAAACA 706

RESULT 4

LOCUS AB022014 696 bp mRNA ROD 09-JAN-1999
DEFINITION Rattus norvegicus mRNA for gankyrin homologue, complete cds.
ACCESSION AB022014
VERSION AB022014.1 GI:4127228
KEYWORDS rat gankyrin homologue.
SOURCE Rattus norvegicus adult female placenta cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 696)

AUTHORS
TITLE
JOURNAL

Higashitsuji, H. and Fujita, J.
Direct Submission
Submitted (04-JAN-1999) to the DDBJ/EMBL/GenBank databases. Hiroaki
Higashitsuji, Faculty of Medicine, Kyoto University, Clinical
Molecular Biology; Shogoin Kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto
606-8507, Japan (E-mail:hhigashi@viru.s.kyoto-u.ac.jp,
Tel:81-75-751-3753, Fax:81-75-751-3750)

REFERENCE

2 (bases 1 to 696)
Higashitsuji, H. and Fujita, J.
Cloning of rat gankyrin homologue containing ankyrin repeats
Published Only in Database (1999) In press
Location/Qualifiers
source
1. .696

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/dev_stage="adult"
/sex="female"
/tissue_type="placenta"
1. .696
/codon_start=1
/product="gankyrin homologue"
/protein_id="BAA36954.1"
/db_xref="GI:4127228"

/translation="MEGCVNLMVNCNLYANGKIDELKESILADKSLATFRDQDSRTAL
HWACSGAGTEIVEFLQLQVFNKDDAGSDPHIAASAGRDEIVKALLIKGAQVNAV
NONGCTALHYAASKNRHEIAVMLLEGGANPDANKHYDATAMHRAAAKGNLKMVHLLIF
YKASTNIQDTGNTPLHLACDERVEEAKLLVQSQASIIYIENKEKTPLOVAKGGLGL
ILKRIAESEASM"

BASE COUNT 211 a 139 c 183 g 163 t

ORIGIN

Query Match 76.0%; Score 593; DB 94; Length 696;
Best Local Similarity 91.9%; Pred. No. 1.6e-169;
Matches 626; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 97 atggagggggtgtgtctaaactaatgtctgcaacctggcctacagcgggaagctgaa 156
DB 1 ATGGAGGGGTGTGTCTAACCTAATGTGTCTAACCTGGCCTACAAACGGGAAGCTGAT 60
QY 157 gattgaaggagagatctctgcccataaaatccctggctactagaaactgaccagcagc 216
DB 61 GAGTTGAAGGAAGCATTTTGGCTCATAGTCTCTGGCCACTAGAACTGATCAGGACAGC 120
QY 217 aqaactgcattgcactgggcgtcagctgcagtcagacacagaaatgttgaaattttgttg 276
DB 121 AGAACAGCATTCCTACTGGGCATGCTCAGCTGGTGCATACAGAAATTTGTAATTTGCTG 180
QY 277 caactggagtgccagtgaaatgataaagacgatgcaggttggtctcctctcatattgcg 336
DB 181 CAACTTGGAGTCCAGTAAATGAAAAGACGATGAGGTGGTCTCTCTCTTCATATTGCT 240
QY 337 gcttctgctggcgggatgagattgtaaaagccctctctggaaaaggtgctcaagtgaat 396
DB 241 GCTTCGCTGGCGGATGAGATTGTAAAGCCCTCTGATAAAGGGGACACAGTGAAT 300
QY 397 gctgtcaatcaaaatggctgtactcccttacattatgcagcttcgaaaaaacagcgatgag 456
DB 301 GCGGTCAATCAGAAATGGCTGTCACGGCCCTTCATTATGACAGTTCACCAAGAAATAGGATGAG 360
QY 457 atcggtctcatattactggaagcgggctcaatccagatgctaaagaccattatgagct 516
DB 361 ATTGCTGTATTGTACTAAGGTGGGGCTAATCCAGATGCTAAGNACCATTTATGATGCT 420
QY 517 acagcaatcacccgggcagcagcccaagggttaactggaagatgattcaatcctctctgtac 576
DB 421 ACAGCAATGCACCGGCAGCAGCAGCAAGGGTAACTTGAAGATGGTTCATATCTCTGTGTC 480
QY 577 tacaagcatccacaacatccaaagacacactgaggggttaaacctcctctacacttagcct 636
DB 481 TACAAAGCATCCCAAAACATCCAAAGATATGAGGGTAAACCTCTCTACACTTAGCTGT 540
QY 637 gatgagagagagtggaagaaagcaaaactgctggtgtgtcccaagagcagattattacatt 696

```

|||||
Db 541 GATGAGGAGAGTGCAGCAAGCAAAATGCTGCTGACCAAGGCAAGTATTACATT 600
|||||
QY 697 gagaaataagaagaaagacacccctgcaagtgccaaagggcgctgggttttaatactc 756
|||||
Db 601 GAAAAAAGAGAGAAAGACACCGCTGCAAGTCGCCAAAGGGGGCCCTGGGTTTAATAC 660
|||||
QY 757 aagagaatggtggaaggttaa 777
|||||
Db 661 AAAAGAATCGCAGAAAGTGAA 681
|||||

RESULT 5
AB022022 696 bp mRNA ROD 09-JAN-1999
LOCUS Mus musculus mRNA for gankyrin, complete cds.
DEFINITION AB022022
ACCESSION AB022022
VERSION AB022022.1 GI:4127259
KEYWORDS gankyrin.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL Higashitsuji, H. and Fujita, J.
Submitted (03-JAN-1999) to the DDBJ/EMBL/GenBank databases. Hiroaki
Higashitsuji, Faculty of Medicine, Kyoto University, Clinical
Molecular Biology; Shogoin Kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto
606-8507, Japan (E-mail:hhigashie@viru1.virus.kyoto-u.ac.jp,
Tel:81-75-751-3753, Fax:81-75-751-3750)
2 (bases 1 to 696)
REFERENCE
AUTHORS Higashitsuji, H. and Fujita, J.
TITLE Cloning of mouse gankyrin containing ankyrin repeats
JOURNAL Published only in DataBase (1999) In press
FEATURES
location/Qualifiers
1..696
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="adult"
/sex="female"
/tissue_type="placenta"
1..696
/codon_start=1
/product="gankyrin"
/protein_id="BAA36969.1"
/db_xref="GI:4127260"
/translation="MEGCVSNIMICNLAYSGLKDELKERILADSKSLATRTDQDSRTAL
HWACSGAHTIEVFELQLQVPVNDKDGWSPLHTAASAGRDEIVKALLVKGHNVSV
NONGCTPLHYAAKSRHETSVMLEGGANPDADKHVDATAMHRAAKGNLKWVHTLIF
YKASNIQDTGENTPLHLACDEERYEAKFLVTQGCASIYIENKEEKTPLQVAKGGLGL
ILKRLAESEEAASM"
BASE COUNT 205 a 142 c 180 g 168 t 1 others
ORIGIN

Query Match 74.9%; Score 584; DB 94; Length 696;
Best Local Similarity 91.0%; Pred. No. 8.8e-167;
Matches 620; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 97 atggagggtgtgtctaacctaatggtctcaacctggctcaacctggctacagcggaagctggaa 156
|||||
Db 1 ATGGAGGGGTGTGTCTAACATATGATCTGAACCTGGCTTACAGTGGGAGCTGGAT 60
|||||
QY 157 gagtgaagagagattgtgcgcgcgataaaatccctggctactagaactgaccaggagcacg 216
|||||
Db 61 GAGTTCAAGGAGCGCATTTTGGCTGATAAATCTCTGGCTACTAGAACTCATCAGGACACC 120
|||||
QY 217 agaactgattgcactggcgatctcagctgacatacacagaattgtgaaattttgttg 276
|||||
Db 121 AGAACAGCTTTTGCACTGGCGATGCTCAGCTGGCCATACAGAAATTTTGAATTTCTGCTG 180
|||||
QY 277 caacttggagtgccagtgaatataaagacgatgcaggtgtgtctctctctcatattg 336
|||||

```

```

|||||
Db 181 CAACITGGAGTGCCACTNAATGATAAAGATGACGCAAGTGTGCTCTCTCTCATATTGCT 240
|||||
QY 337 gcttctgctggccgggagtgagattgtaaaagccctctctgggaaaaagtgctcaaggaat 396
|||||
Db 241 GCCTCGCGCTGGCGGGATGAGATTGTAAAGGCCCTTCTGTGTGAAGAGTGCACATGTTAAT 300
|||||
QY 397 gctgtcaatacaaaatggctgtactcccttacattatgcagcttcgaaaaacagggcatgag 456
|||||
Db 301 TCTGTCAATCAAAACGGCTGCACCTCCACTCCATTATGACAGCTTCGGAAGAATAGGCATGAG 360
|||||
QY 457 atcgtgtcatgttactggaagcgggggttaataccagatgctaaagaccattatgaggt 516
|||||
Db 361 ATTCTCTGTTATTGTTACTAGCAAGTGGGGCTAACCCAGATGCGAAGCACCATTTACGATGCT 420
|||||
QY 517 acagcaatgcacggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
|||||
Db 421 ACAGCAATGCACCGGCGCAGCAGCAAGGTAATCTTGAAGATGCTTCACATCTCTTCTGTTTC 480
|||||
QY 577 tacaagcatcacaaacatccaagacactgagggtaacacactctctacacttagcctgt 636
|||||
Db 481 TACAAGCAATCCACAAACATCCCAAGACACTGAGGTTAACACTCTCTACACTTAGCCTGT 540
|||||
QY 637 gatggagagagtggtggaagaagcaaaactgctggtgtcccaaggagcaagattattacatt 696
|||||
Db 541 GATGAAGACAGAGTGGAAGAGGCAAAATTTCTGCTGACTCAAGGACCAAGTATTACATT 600
|||||
QY 697 gagaataagaagaaagacacccctgcaagtgccaaagggcgctgggttttaatactc 756
|||||
Db 601 GAGATAAAGAGAAAGAACACCCCTGCAAGTTGCCAAAGGGGGCCCTGGGTTTAATAC 660
|||||
QY 757 aagagaatggtggaaggttaa 777
|||||
Db 661 AAGAGACTAGCAGAAAGTGAA 681
|||||

RESULT 6
AC012149 180905 bp DNA HTG 06-SEP-2000
LOCUS Homo sapiens chromosome 3 clone RP11-119E13, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.
ACCESSION AC012149
VERSION AC012149.8 GI:9438824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muzny, D.M., Adams, C., Bailey, M., Barberia, J., Blankenburg, K.,
Bodota, B., Bouck, J., Bowie, S., Brooks, A., Buhay, C., Bunac, C.,
Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,
David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,
Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,
Forcum-Tansey, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,
Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hogues, M.,
Holloway, C., Hosak, H., Jackson, L.B., Jackson, L., Jia, Y., Jones, M.,
Kelly, S., Kondejewski, N., Kong, Y., Kovar, C., Leal, B., Li, Z.,
Lichtarge, O., Liu, J., Liu, W., Logan, O., Lozano, R.J., Lu, J.,
Lucier, R., Martin, R., Martinez, C., McLeod, M.P., Mei, G., Morgan, M.,
Morris, S., Nash, S., Nelson, A., Nguyen, R., Nguyen, N., Nguyen, S.,
Oswal, G., Parish, B., Paxton, S., Payton, B., Perez, L., Pu, L.L.,
Quiles, M., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S.,
Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Sugeng, R.,
Tabor, P., Taylor, T., Vasquez, L., Vinson, R., Vo, Q., Wahbah, M.,
Watlington, S., Weinstock, G., Weinstock, I.R., Williamson, A.,
Worley, K., Wren, J., Wrensford, G., Yu, W., Zhou, X., Nelson, D. and
Gibbs, R.
Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 180905)
AUTHORS Worley, K.C.
TITLE Direct Submission

```


Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>

RP5-914P20 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP5-914P20 The true left end of clone RP5-955M13 is at 37950 in this sequence.

FEATURES

source	Location/Qualifiers	
	1..105334	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="20"	
	/map="q13.13-13.2"	
	/clone="rp5-914p20"	
	/clone_lib="RPCI-5"	
repeat_region	1..212	
	/note="AluJo repeat: matches 35..237 of consensus"	
repeat_region	682..985	
	/note="AluX repeat: matches 1..304 of consensus"	
repeat_region	1184..1354	
	/note="MIR repeat: matches 17..196 of consensus"	
repeat_region	1932..2066	
	/note="FLAM_C repeat: matches 1..131 of consensus"	
repeat_region	2292..2398	
	/note="L1ME3 repeat: matches 4915..5009 of consensus"	
repeat_region	2399..2692	
	/note="AluJo repeat: matches 1..298 of consensus"	
repeat_region	2693..3111	
	/note="L1ME3 repeat: matches 5009..5354 of consensus"	
repeat_region	3112..3407	
	/note="AluX repeat: matches 1..297 of consensus"	
repeat_region	3408..3456	
	/note="L1ME3 repeat: matches 5354..5396 of consensus"	
repeat_region	3506..3641	
	/note="AluJb repeat: matches 163..294 of consensus"	
repeat_region	3662..3948	
	/note="L1ME3 repeat: matches 5612..5929 of consensus"	
repeat_region	4569..4869	
	/note="L1M4 repeat: matches 3660..3981 of consensus"	
repeat_region	5879..5988	
	/note="MER5B repeat: matches 64..177 of consensus"	
repeat_region	6820..7120	
	/note="AluSp repeat: matches 1..297 of consensus"	
repeat_region	7311..7457	
	/note="AluJb repeat: matches 141..288 of consensus"	
repeat_region	8386..8684	
	/note="AluSg repeat: matches 1..301 of consensus"	
misc_feature	complement(8685..9109)	
	/note="match: GSS: Em:AQ151284"	
repeat_region	9525..9838	
	/note="AluY repeat: matches 5..302 of consensus"	
misc_feature	complement(10462..10988)	
	/note="match: GSS: Em:B89004"	

misc_feature	10509..10953	
	/note="match: STS: Em:H18297"	
misc_feature	complement(11204..11523)	
	/note="match: STS: Em:F06436"	
misc_feature	complement(11209..11577)	
	/note="match: STS: Em:H79769"	
repeat_region	11579..11614	
	/note="18 copies 2 mer aa 97% conserved"	
misc_feature	complement(join(11790..15959,23464..23556,25336..25448,50098..50272,52273..52422))	
	/gene="dJ914P20.2"	
	/note="match: ESTs: Em:BE112660 Em:BE103176 Em:BF464091 Em:BS853100 Em:AW139427 Em:BF445957 Em:AW173331 Em:AA111726 Em:BF318209 Em:AW389690 Em:BF523190"	
	/evidence="not experimental"	
	/product="dJ914P20.2.2 (activity-dependent neuroprotective protein (KIAA0784, ADNP), isoform 2 (alternative 5' UTR))"	
	11790..11790	
	/note="match: STS: Em:G30243"	
misc_feature	complement(11790)	
	/gene="dJ914P20.2"	
polyA_site	complement(join(11790..15959,23464..23556,25336..25754))	
	/gene="match: cDNAs: Em:L25107 Em:AB020670 Em:AF068198 Em:AB018327	
	match: ESTs: Em:AI600029 Em:AA183241 Em:AA275189 Em:AA745620 Em:AA830325 Em:AI020367 Em:AA061696 Em:AI243490 Em:AI146218 Em:AV016003 Em:AA259872 Em:AI590352 Em:AA227896 Em:C14463 Em:244748 Em:AI071938 Em:AA874881 Em:AA495739 Em:D81665 Em:AI277858 Em:T34137 Em:AA371316 Em:AA067127 Em:R13866 Em:W32680 Em:AA067128 Em:W12764 Em:W45435 Em:AA528083 Em:AA373256 Em:AA249727 Em:AA790780 Em:243469 Em:AA636938 Em:C02493 Em:AI078094 Em:AI653383 Em:C15608 Em:AI605766 Em:AI636691 Em:R20042 Em:AA111726 Em:AV014718 Em:AA285030 Em:H26842 Em:AI548767 Em:AI394588 Em:R61133 Em:R23630 Em:AA207035 Em:AA525060 Em:AA207036 Em:AA600642 Em:AA243373 Em:AA739201 Em:AA291097 Em:AI007177 Em:T86563 Em:AI181625 Em:AI715802 Em:AA227596 Em:T84549 Em:AA874149 Em:AI414826 Em:AA863656 Em:AI718275 Em:AI555864 Em:D44915 Em:AA354206 Em:AA722822 Em:AA495802 Em:AA806934 Em:W41625 Em:AI692233 Em:AI050875 Em:DI2086 Em:AA281988 Em:AI298135 Em:L44402 Em:AI655187 Em:AI674538 Em:AI548629 Em:AI548468 Em:T10165 Em:W40213"	
	/evidence="not experimental"	
	/product="dJ914P20.2.1 (activity-dependent neuroprotective protein (KIAA0784, ADNP))"	
gene	complement(join(11790..15959,23464..23556,25336..25754,50098..50272,52273..52422))	
	/gene="dJ914P20.2"	
	11791..12219	
misc_feature	/note="match: STS: Em:T86741"	
misc_feature	11796..12023	
	/note="match: STS: Em:T10164"	
polyA_site	complement(11796)	
	/gene="dJ914P20.2"	
polyA_signal	complement(11814..11819)	
	/gene="dJ914P20.2"	
CDS	complement(join(12852..15959,23464..23556,25336..25443))	
	/gene="dJ914P20.2"	
	/note="match: cDNAs: Em:AB018327"	
	/codon_start=1	
	/evidence="not experimental"	
	/product="dJ914P20.2.1 (activity-dependent neuroprotective protein (KIAA0784, ADNP))"	
	/protein_id="CAB53748.2"	
	/db_xref="GI:13276226"	
	/db_xref="SPTREMBL:O94881"	
	/translation="MFOLPVNVLGSLRKARKTKKILSDIGLEYCKEHIEDKPOEPNDFYLNKNTWEDVGLWDPSTKNDYRTKPCSCACPFSSKPFSAVKSHFNVKSDFE NRILLNCPYCFNADKDKLETHIKIFHAPNAPSPPSSLSLTFDKXKNDGLPKQNDVSOAVYKCKTFRDPLYEIVRKHYIREFHQHVAAPYIAKAGEKSLNGLPKQNSVSSIHCKRCLFMPKQIVHVEDHRIQGYVTAMIGHTNVVVRSPKPLIAPKPQDKKSMGLPPIRGLASGNVRSLSQOQVNRSLIPKPNLNLSTGVNMSSVHLQQNNYGV"	

```

KSVGGYVGVGOSMRGLGNNAPVSTPOOSVSKVOLLPSNGRSYGLGSEORSOAPARY
SIQSANASLSGQKSPSLQSOASRVLGSSSKPAAATGPPGNTSSSTQKWKICT
ICNELFPENYVSHFEKEHKAEMRVAVANYIMKHNETSKCLYCNRLYPTDNLNHL
IHGLSCPYCRSTFNDVDEKMAHMRVHIIDEEMPKTSTLSFDLTLOQGSHTNLHLV
TTYNEURDAPASVYHAQNNPPVPKQPKQVKEKADIPKSSPOAAYPYKDKVGKTLIC
PLCFSLILKGPISDALAHLRHRHVIQVHPVEKRLTKYKICHLGVAVTSSMTASTITL
HIVHCRGVGKQNGODKTNAPSLNOSPSPAPVARTYBOMEFFPLKKRKLDDSDSPS
FIEEKPEPVLALDPKDHEDDSYARKSFLTKYFNKOPYPTRELEKLAASLWMS
DIASHFSNKRKVCVRCKEYKYPVLLGFNMKELNKKVKHEMDFDAEWLFENHDEKDSRV
NASKTADKLNGLKEDDSSDSFENLEESNESGSPFPDVEFPKISNDNPEHVLK
VIPEDASEEKLKDGDSKYETIHLTEPTKLMHNASDSEVDQDDVNDWKGASPS
ESGPGSOQVSPEDNTCMKPGTWSDESSDARSKPAAKKTAKMGDQREQLKWN
SSYGVGEFWSKDQSQWKNASENDERLSNPQLEWQNSTIDSEDEGEQFDNMTDGVAEPM
HSLAGVXLSLSSQA"
misc_feature
complement(13731..14156)
/gene="d0914P20.2"
/note="match: STS: Em:G37216"
13814..14173
/note="match: STS: Em:L44402"
16060..16350
/note="AluSg repeat: matches 1..291 of consensus"
16607..17008
/note="Tigger3(Golem) repeat: matches 2649..3028 of
consensus"
17009..17274
/note="AluJo repeat: matches 1..286 of consensus"
17275..17389
/note="Tigger3(Golem) repeat: matches 2542..2649 of
consensus"
Query Match 68.2%; Score 532; DB 92; Length 105334;
Best Local Similarity 89.9%; Pred. No. 1.5e-150;
Matches 629; Conservative 0; Mismatches 60; Indels 11; Gaps 5;

Qy 82 tgcctggagacagcgaatgagggggtgtgtcttaacctaatgctgcgaacctggcctac 141
Db 35160 TCCTGGGAAAGCAACAGACGGGTGTCTCTAAACTTATGTCTGTCAACTTGGCCTAC 35219

Qy 142 agcgggaagctggaagagtgatctgtgcgcataataatccctggctactaga 201
Db 35220 AGAGGGAAGCTGGAGAGGTGAAGGAGAGATCTGGCCATAAACCCTGGCTATGAGA 35279

Qy 202 actgaccagacagcagaatgcattgctgctggcgcagctgcagctgacatacagaat 261
Db 35280 ACTGACCAGGA---CAGAAGTGCAATTCATTTGGCAGTCTCAGCCATACATACATAAT 35336

Qy 262 gttgaattttgtgcaactgagtgccagtggaatgataaagacgatgc-aggttgctc 320
Db 35337 ATTGAATTTTGTTCACCTGGAGTGCCAGTGAATGATAAAGACTATGCGAGATGGTGC 35396

Qy 321 tcctcttcatttgctgtctgctgcccgggatgagatgtgaaagccctcttgggaaa 380
Db 35397 TCCCTTCAGACTGCTGTTGTGTCGCCAGATGAGACTGTAAAGCCCTCTGGGAAA 35456

Qy 381 aggtgctcaagtgaatgctgtcaatacaaatggctgtactcccttaccattatgcagcttc 440
Db 35457 AGGTGCTCAAGTGAATGCTGTCAATCAAAATGGCTCTACTCCCTACATCATGAGCTTC 35516

Qy 441 gaaaaaaggagatgagatgcgtgtctgtactggaagggcggggctaatccagatgctaa 500
Db 35517 CAAAACAGGATGAGATGTGCTCTCATGTTACTAGAGAGGTGGGCTAATCCAGATGTAA 35576

Qy 501 ggaccattatgagctacagcaatgacccggcgagccagcaagggtaacttgaagatgat 560
Db 35577 GGATCATTTATGAGGCTACAGCAAGCACCAGCCACGCAAGGGAATTCATCAAGATGAT 35636

Qy 561 tcatctctctgtactacaagcatccacaacatccagacactgaggggtaacatcc 620
Db 35637 TCATATCTCTGTACTACAAGCATCCACAATCATCCAGACACTGAGGGTAAACATCC 35696

Qy 621 tctaacactagccttgatgaggaagtggaaggaagcaaaactgctggtcccaagg 680
Db 35697 TCCACACTTAGTCTGTGAT-----AAGAGTGAAGAGCAAACTGCTGTGTCCTCCCAAGG 35751

```

```

Qy 681 agcaagtattacattgagaataaagaagaacacacccctgcaagtggccaaagtgg 740
Db 35752 ACCAA-TATTACATAGAGATAAAGAAGAAAGAACCC-ACAACTGGCCAAAGGTGC 35809

Qy 741 cctgggttatactacatcaagaagaatgggtggaagggttaaca 780
Db 35810 CCTGGTTTAGTACTCAAGAGAATGGTAGAAGGTTAAACA 35849

RESULT 8
AC009863/c
LOCUS AC009863 175550 bp DNA HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 3 clone RP11-573D15 map 3, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
AC009863
AC009863.3 GI:8096876
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175550)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 3, clone RP11-573D15
TITLE Unpublished
AUTHORS 2 (bases 1 to 175550)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (03-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced gi:6514025.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1835
Center clone name: 573_D15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 139143 bases at least Q40
Consensus quality: 158720 bases at least Q30
Consensus quality: 168324 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 174050; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```


During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Embl:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human

chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr13>

RP11-505F3 is from the library RPI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pBACE3.6

This sequence is the entire insert of clone RP11-505F3.

FEATURES

source

1. 199882
/location=Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-505F3"
/clone_lib="RPI-11.2"

BASE COUNT 5887 a 40074 c 40071 g 60866 t

ORIGIN

Query Match 51.1%; Score 398.2; DB 90; Length 199882;
Best Local Similarity 93.6%; Pred. No. 9.8e-110;
Matches 437; Conservative 0; Mismatches 28; Indels 2; Gaps 2;
QY 79 agttgtggagcagcaaaatggaggggtgtgtgtacacctaattgtgtcgaacctggcc 138
Db 101505 AGTTCTCTGGAAAGCAAAATGGAGGGGTGTGTGTCTAACTTAATGCTGCAATCTGCC 101564
QY 139 tacagcgggaagcgtggaagattgaagagagattcttgcgcgataaatccctggctact 198
Db 101565 TACAGCAGGAAGCTGGAAAGATTCAAGGAGGGTATTCTGGCGGATAAATCCCTAGCTACT 101624
QY 199 agaactgaccagcagcagcaactgcattgcactggcgatctcagctggacacacagaa 258
Db 101625 AGAAGTACACAGGACAGCAGACAGCATTGCTCAGTGGCATGCTCAGTGGACATAAAGGA 101684
QY 259 attgttgaaattttgttgcacattggagtcgagtgaaatgataaagacgatcgaggttgg 318
Db 101685 ATTGATGAATTTTGTTCACATTTGGAGTGGCAGTGAATGATAAAGACGATGCAAGTTGG 101744
QY 319 tctctcttcattatgcgcttctgcgcggagatagattgataaagcccttcaggga 378
Db 101745 TCTCTCTC-TCAAAATGCGGGCTTCTGCTGCGCGGATGAGATTGATAAAGCCCTCTCGGA 101803
QY 379 aaagtgctcgaagtgaatgctgtcaatcaaaatggctgtacttcccttaccattatgcagct 438
Db 101804 AAAGTGCTCAGTGAATGCTGCTCAATCAAAATGACTTACTCCCTACATATGACGCT 101863
QY 439 tcgaataaacagcagcagatagatcgtctgactgtaagcggcgagcgttaacccagatgct 498
Db 101864 TCCAAAATAGGCAATGAGATTGCTGCTGATGCTACTAGAGGCGGGCTTAATCCTCATGCT 101923
QY 499 aaggaccattatgagcgtacagcaatgacccggcgagcagcagcaagg 545
Db 101924 AAGGA-CATTATGAGGCTACAGCATTGCTACTAGGAGGAGCCAGGG 101969

RESULT 10

AC072018

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC072018 145483 bp DNA HTG 06-SEP-2000
Homo sapiens chromosome 3 clone RP11-24009, *** SEQUENCING IN
PROGRESS ***, 40 unordered pieces.

AC072018

AC072018.3 GI:9966538

HTG: HTGS_PHASE1.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145483)

Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
Bodota,B., Bouck,J.J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,R.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wahbah,M.,
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 145483)

Worley,K.C.

Direct Submission

Submitted (07-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Sep 5, 2000 this sequence version replaced gi:8572872.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HAHU

Center clone name: RP11-24009

----- Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye; 93% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 112126 bases at least Q40

Consensus quality: 127972 bases at least Q30

Consensus quality: 136731 bases at least Q20

Estimated insert size: 134186; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 40 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 10264: contig of 10264 bp in length

* 10265 10364: gap of unknown length
* 10365 contig of 7422 bp in length
* 17786 17886: gap of unknown length
* 17787 24658: contig of 6772 bp in length
* 24659 24758: gap of unknown length
* 24759 30985: contig of 6227 bp in length
* 30986 31085: gap of unknown length
* 31086 36928: contig of 5843 bp in length
* 36929 37028: gap of unknown length
* 37029 43546: contig of 6518 bp in length
* 43547 43646: gap of unknown length
* 43647 50051: contig of 6405 bp in length
* 50052 50151: gap of unknown length
* 50152 55816: contig of 5665 bp in length
* 55817 55916: gap of unknown length
* 55917 62446: contig of 6530 bp in length
* 62447 62546: gap of unknown length
* 62547 67354: contig of 4808 bp in length
* 67355 67454: gap of unknown length
* 67455 72321: contig of 4867 bp in length
* 72322 72421: gap of unknown length
* 72422 76988: contig of 4567 bp in length
* 76989 80703: gap of unknown length
* 80704 80803: gap of unknown length
* 80804 84396: contig of 3593 bp in length
* 84397 84496: gap of unknown length
* 84497 88793: contig of 4297 bp in length
* 88794 88893: gap of unknown length
* 88894 92423: contig of 3530 bp in length
* 92424 92523: gap of unknown length
* 92524 95759: contig of 3236 bp in length
* 95760 95859: gap of unknown length
* 95860 99818: contig of 3959 bp in length
* 99819 99918: gap of unknown length
* 99919 103132: contig of 3214 bp in length
* 103133 103232: gap of unknown length
* 103233 105760: contig of 2528 bp in length
* 105761 105860: gap of unknown length
* 105861 108187: contig of 2327 bp in length
* 108188 108287: gap of unknown length
* 108288 111981: contig of 3694 bp in length
* 111982 112081: gap of unknown length
* 112082 114420: contig of 2339 bp in length
* 114421 114520: gap of unknown length
* 114521 116992: contig of 2472 bp in length
* 116993 119463: contig of 2371 bp in length
* 119464 119563: gap of unknown length
* 119564 121783: contig of 2220 bp in length
* 121784 121883: gap of unknown length
* 121884 123681: contig of 1798 bp in length
* 123682 123781: gap of unknown length
* 123782 126198: contig of 2417 bp in length
* 126199 126298: gap of unknown length
* 126299 128556: contig of 2258 bp in length
* 128557 128656: gap of unknown length
* 128657 129817: contig of 1161 bp in length
* 129818 129917: gap of unknown length
* 129918 131742: contig of 1825 bp in length
* 131743 131842: gap of unknown length
* 131843 133420: contig of 1578 bp in length
* 133421 133520: gap of unknown length
* 133521 134903: contig of 1383 bp in length
* 134904 136266: contig of 1263 bp in length
* 136267 136366: gap of unknown length
* 136367 137771: contig of 1405 bp in length
* 137772 137871: gap of unknown length
* 137872 138996: contig of 1125 bp in length
* 138997 139097: gap of unknown length
* 139097 140945: contig of 1849 bp in length
* 140946 141045: gap of unknown length

* 141046 142678: contig of 1633 bp in length
* 142679 142778: gap of unknown length
* 142779 144032: contig of 1254 bp in length
* 144033 144132: gap of unknown length
* 144133 145483: contig of 1351 bp in length.
FEATURES
Location/Qualifiers
1..145483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-24009"
BASE COUNT 38303 a 31328 c 31243 g 40572 t 4037 others
ORIGIN
Query Match 28.7%; Score 223.8; DB 74; Length 145483;
Best Local Similarity 91.9%; Pred. No. 1.3e-56;
Matches 249; Conservative 0; Mismatches 17; Indels 5; Gaps 1;
QY 510 tgagctacagcaatgcaccggcgagcagccagggtaacttgaagatgattcatatcct 569
|||||
Db 119572 TGAGGCTACAGCAATGCACCGGTACAGCAGCCAGGGTAACCTTGAGATGATTCCTATCCT 119631
|||||
QY 570 tctgtactacaagaatccacaacatccaaagacactgagggtaaacactcctctacactt 629
|||||
Db 119632 TCTATACTACAGCAATCCACGACATCCAGACACTGAGGGTAACACTCTCTACACTT 119691
|||||
QY 630 agcctgtgatgaggagagagtggagaagcaaaactgctggtgctcccaaggagcaagtat 689
|||||
Db 119692 ACCCTGTGATGAGGAGAGAGTGGAAAGGA-----GCAAGTGTCCCAAGGAGCAAGTAT 119746
|||||
QY 690 ttacattgagaataagaagaacacacccctgcaagtgcgaaggctggctgggttt 749
|||||
Db 119747 TCACATGGAGAATAAGAAAGAAAGACACCCCTGCAAGCGGCAGAGGTGGCTACGTTT 119806
|||||
QY 750 aatactcaagaagaatggggaaggttaaaac 780
|||||
Db 119807 AATACTCAAGTCAATGGTGAAGGTTAAACA 119837
|||||
RESULT 11
HS889N15/c
LOCUS HS889N15 123395 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from clone 889N15 on chromosome Xq22.1-22.3.
Contains part of the gene for a novel protein similar to X.laevis
Cortical Thymocyte Marker CTX, the possibly alternatively spliced
gene for 26S Proteasome subunit p28 (Ankyrin repeat protein), a
novel gene and exons 36 through 45 of the COL4A6 for Collagen Alpha
6(IV). Contains ESTs, STSS, GSSs and a putative CpG island,
complete sequence.
ACCESSION AL031177
VERSION AL031177.1 GI:4071056
KEYWORDS HTG; 26S Proteasome subunit p28; Ankyrin repeat; COL4A6; Collagen
Alpha 6(IV); Cortical Thymocyte Marker; CpG island; CTX.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 123395)
AUTHORS Grafham, D.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1999) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
COMMENT On Dec 29, 1998 this sequence version replaced gi:4034456.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence is the entire insert of clone 889N15. This sequence
has been finished according to sequence map criteria as follows. An


```
/note="match: GSS B42534"
36027. .36306
/note="AluX repeat: matches 1. .286 of consensus"
complement(join(37183. .38105,40664. .40830,40931. .41077,
41730. .41828,44385. .44596))
/gene="dJ889N15.2"
/note="match: CDNAS DB3137 AB009619 E12804
match: ESTs AA513142 W31009 AA491647 241316 R15438 N98727
AA156429 AA460532 A1274952 BA810950 A1215759 A1017040
AA495954 A1201389 AA861216 N26039 N80914 AA916950 N52252
A1268842 A1449036 AA700682 A149035 R77104 H04970 AA210845
AA203428 T98190 NA0402 AA005090 AA025485 N91106 AA035825
R83144 R10433 AA495778 AA957491 H47845 AA608308 H10670
AA271098 AA147557 AA425037 H25881 AA050244 AA460531 F08311
RG0087 W78023 AA385621 R13947 R13250 AA146798 N56112
T30546"
/revidence=not_experimental
/product="dJ889N15.2.1 (26S Proteasome subunit p28
(Ankyrin repeat protein)) (isoform 1)"
complement(37183. .44596)
/gene="dJ889N15.2"
complement(37197. .37202)
/gene="dJ889N15.2"
polyA_signal

Query Match 27.1%; Score 211; DB 92; Length 123395;
Best Local Similarity 100.0%; Pred. No. 1e-52;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggtagctctaacgctgtttttagctggcgtagcggagcggcgagcgtgagcgggc 60
Db 44594 TGGTAGGCTCTAACGCTGTTTGTACTGGCGTAGCGGAGCGCGGCGGCGG 44535

QY 61 gttgctgcgcgcacaagttagtctgtgggacagcgaatgaggggtgtgtctaaccta 120
Db 44534 GTTGCTCGCGCGCAAGTAGTGTCTCGGACAGCGNAATGAGGGGTGTGTCTAACCTA 44475

QY 121 atggtctcaaacctggcctacagcggaagctggaagtgtagagtgagagattctggcc 180
Db 44474 ATGCTGTGCAACCTGGCCTACACGGGAGCTGGAAGAGTTGAAGGAGAGTATTCTGGCC 44415

QY 181 gataaatccctggctactgaactgaccagg 211
Db 44414 GATAAATCCTGGCTACTAGAACTGACCAGG 44384

RESULT 12
AL137065/c
LOCUS
DEFINITION Homo sapiens chromosome X clone RP11-4P13 map q22.1-23, ***
SEQUENCING IN PROGRESS ***, 5 unordered pieces.
ACCESSION AL137065
VERSION AL137065.6 GI:13567838
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178492)
Direct Submission
Submitted (07-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SN, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: cloner@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:12750785.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bA4P13
----- Summary Statistics
Assembly program: XGAP4; version 4.5
```

```
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ET-amerham; 5% of reads Chemistry:
Dye-terminator Big Dye; 94% of reads
Consensus quality: 177346 bases at least Q40
Consensus quality: 177803 bases at least Q30
Consensus quality: 177954 bases at least Q20
Insert size: 178092; sum-of-contigs
Insert size: 162815; 5.8% error; agarose-fp
Quality coverage: 6.98x in Q20 bases; sum-of-contigs Quality
coverage: 7.66x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 6091: contig of 6091 bp in length
* 6092 6191: gap of 100 bp
* 6192 35640: contig of 29449 bp in length
* 35641 35740: gap of 100 bp
* 35741 49648: contig of 13908 bp in length
* 49649 49748: gap of 100 bp
* 49749 116198: contig of 66450 bp in length
* 116199 116298: gap of 100 bp
* 116299 178492: contig of 62194 bp in length.
FEATURES
Location/Qualifiers
source
1..178492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22.1-23"
/clone="RP11-4P13"
/clone_lib="RPC1-11.1"
1..6091
/note="assembly fragment:01906
fragment_chain:1"
6192..35640
/note="assembly fragment:03330
fragment_chain:1"
35741..49648
/note="assembly fragment:01670
fragment_chain:1"
49749..116198
/note="assembly fragment:01340"
116299..178492
/note="assembly fragment:01657"
BASE COUNT 53840 a 35194 c 34622 g 54436 t 400 others
ORIGIN

Query Match 27.1%; Score 211; DB 78; Length 178492;
Best Local Similarity 100.0%; Pred. No. 1e-52;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggtagctctaacgctgtttttagctggcgtagcggagcggcgagcgtgagcgggc 60
Db 100941 TGGTAGGCTCTAACGCTGTTTGTACTGGCGTAGCGGAGCGCGGCGGCGG 100882

QY 61 gttgctgcgcgcacaagttagtctgtgggacagcgaatgaggggtgtgtctaaccta 120
Db 100881 GTTGCTCGCGCGCAAGTAGTGTCTGGGACAGCGNAATGAGGGGTGTGTCTAACCTA 100822

QY 121 atggtctcaaacctggcctacagcggaagctggaagtgtagagtgagagattctggcc 180
Db 100821 ATGCTGTGCAACCTGGCCTACACGGGAGCTGGAAGAGTTGAAGGAGAGTATTCTGGCC 100762

QY 181 gataaatccctggctactgaactgaccagg 211
Db 100761 GATAAATCCTGGCTACTAGAACTGACCAGG 100731
```



```
gene
complement(<8359. >16917)
/gene="ci"
/note="CG2125"
/map="101f1-102a1"
/db_xref="FLYBASE:FBan0002125"
/db_xref="FLYBASE:FBgn0004859"
complement(join(8645. .1113,11167. .11331,11396. .11908,
11965. .12605,12668. .13012,16484. .16526))
/gene="ci"
/note="ci" gene product; Nucleotide sequence of the Celera
sequence differs from the published sequence for this
transcript
/codon_start=1
/db_xref="FLYBASE:FBan0002125"
/db_xref="FLYBASE:FBgn0004859"
/protein_id="AAF59373.1"
/db_xref="GI:7304342"
/translat="MKWPTRYLHIFLPSRRAAAVAAATVLPSPCINQHPTDVS
SSVTPSIPTGTSDSIKTSIOPICNENTILGNAGHOHNQPHQVHINVTGPHD
FHPAYRIPGMQLXSLQNTSASFDPYVNCASAFHLAGLGLGSAFLSGRLSSL
GELNAAVAAAAGSLASDFHESVDGNRLGSPRPGGSIIRASLSKRALSSPSYD
SDINSMIRFSPNSLATIMNGSGSAASGSIHSATVAFNPMHSHVSTRLOQIAHL
LRASGLIMPMPQQAASGFSIGHMPTASLRVNDVHNLSDSHIQTITPTVNDV
SOVPAFAFLKLNLDAREKKGFDPVPEQPSSTGGVAQVEADSSOLSQRXNNV
VNNITGPGDKVNSRLDEYINCISIPSEYDCANADTTDKDEPGDFIETNCHWR
SCRIETIDOLEVKHINNDHIOTNKFAVCRWEDCTRGEKPFKAQVLMVHMRHTGE
KPHKCFEGCFKAYSLENLKTHLRSHTEKPYTCYPCGSKAFNASDRKHQNRTH
SNEKPYICAPGCKTKEYPSSLRKHVTVHGAEEYANKKHKGLPLNDANSRLQONS
RHNQHNIDSSPCSDSHLGRKMGHSSFSIKESDISSNHLHLYNGVRASLSLITS
PDLAENLNDWNCDDVDVADLPVLRAMVINGNASASTIGGSVALRQRFGR
LOTGINSITIMLCNTPENRTFGISELNORITELMEPGTDAETIKPILNPTTIGGY
TEDPLQNTSFRNTVSKQSTGSGTQGOFRDQNSTASTYYSQMSRRSSQSVSD
SIPTRPNPSCNSTAFEDYIPSPGSRSSOMSGANCNSFTSTGLPLVKNESKSL
NACINKPNVGVOGVLVNSLPPSPSSHLIATNLKLRKDSFHNFTSGRFSVPSY
HSUHIKNNKPVGENEDKAIASNARQTDVFNINLDPLTNISRESTTTPHSDINWGK
TNNIASINKNDKOLFVTSIKADMAMTSDQHPNERNILDEVEELILPEMLQYLN
VKDTHLNEKHQAVPGVNSSETIASHYREOSNIYTNKQILPLPPSNVDIOPNTTK
FTVQDFAMTAVGGSFQRELSLTPVNEHGHAKCESFHQSQKYNMTDIGKSQAL
PSAHOKEKSNYNOIIDSMTSLPDLNVDISYPRNETENIFKVDHNDNEIQCGIS
QSQMSPTNLNDQSFSTVMOPITTSKLPPEPOKIVGDTQASNTSVAMHDTYQRTL
EYVQSCQNWETNISTNQISQLPGMPVNTLFPDVSSTHPVHGTNMVINDMTTSLT
SLLEENRYLOMMQ"
complement(join(26773. .27392,27494. .27890))
/gene="Rp3A"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT7064"
/db_xref="FLYBASE:FBan0002168"
/db_xref="FLYBASE:FBgn0017545"
complement(<26773. >27890)
/gene="Rp3A"
/note="CG2168"
/map="102A1-102A3"
/db_xref="FLYBASE:FBan0002168"
/db_xref="FLYBASE:FBgn0017545"
complement(join(26927. .27392,27494. .27834))
/gene="Rp3A"
/note="Rp3A" gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript
/codon_start=1
/db_xref="FLYBASE:FBan0002168"
/db_xref="FLYBASE:FBgn0017545"
/protein_id="AAF59372.1"
/db_xref="GI:7304341"
/translat="MAVGKNKGLSKGKGGKKVVDVDFSEKDWYDVKAPNMFTQRI
GKTLVNRQRIASDYLYKGRFEVSLADLQKDDIPERSFKRFLAEDVQDRNLVLCN
FHGMDLTDRQYRSMVKWQTLIEAVEAKTVGYLRFVCFGTAKDQOOSQRTCYAQ
QSOVRKTRARMTDITNEVSGADLKQLVNLALDLSIAKDIKSCQRIYPLHVDYIRKV
KVLKKPRFDVSKLLELGGGKSVKVEAVSSEGAVIDRPEGEPPVQEV"
join(30047. .30635,32975. .32989,33058. .33248,33471. .33516,
35722. .35825,43524. .43632,46598. .46685,46739. .46819,
47903. .48052,55309. .55501,55723. .55890,60156. .60226,
60154. .69555)
/gene="pan"
/product="CT40034"
/db_xref="FLYBASE:FBan0017964"
/db_xref="FLYBASE:FBgn0019664"
<30047. >69555
/gene="pan"
/note="CG17964"
/map="102A1-102A3"
/db_xref="FLYBASE:FBan0017964"
/db_xref="FLYBASE:FBgn0019664"
join(30382. .30503,32975. .33248,33471. .33516,35722. .35825,
43524. .43632,46598. .46685,46739. .46819,47903. .48052,
55309. .55501,55723. .55890,60156. .60226,68154. .69555)
/gene="pan"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
Query Match 14.8%; Score 115.6; DB 5; Length 114334;
Best Local Similarity 51.1%; Pred. No. 1.2e-23;
Matches 271; Conservative 0; Mismatches 259; Indels 0; Gaps 0;
QY 202 actgaccagacagacgaactcattgcaactggcattgctcagctgagacacagaaat 261
Db 82504 ATTGCCCGTGAAGGCCAAACGCCACTTCATGTAGCTTCGTAGTAGGAATTAATATA 82445
QY 262 gttgaattttgttgcacttgagtcgcagtggaatgataaagacagatgcaggttgct 321
Db 82444 ATTATGTTATTCTGCAGCATGGAGCAGAGATAAATGCTCAATCTAATGACAAATATT 82385
QY 322 cctctcatattgctgcttctgctggccgggatgagatgtaaaagcccttctgggaaa 381
Db 82384 GCCTGTCATATTGCTGCAAAAGAACAGAGAAATATAGTCCAAAGTGTACTGGAAAT 82325
QY 382 gttgctcaagtgaatgctgcaatcaaatggctgactcccttaccattatgcagctcg 441
Db 82324 GGTGCCGAAAAAATGCTGTACCAAAAAGGGTTTACTCCACTGCATTTAGCTTCAAA 82265
QY 442 aaaaacaggcatgagatcgctgtctactgttactggaagcggggcgaatccacagatgcta 501
Db 82264 TACGGAAAGCAGAATGTGGTTCAAAATATTACTACAAAATGGAGCTTCTATTGATTTTCAG 82205
QY 502 gaccattatgaggtcacgcaatgcaccggcgagcagcaagggttaactgaagatgatt 561
Db 82204 GGGAAAAATGATGTGACCCCATTCACAGCTGGCCACGCACTACAAATATCCCTCAATAGTA 82145
QY 562 catatccttctgtactacaaagcatccacaacatccaaagacactgagggtaacactcct 621
Db 82144 GAATCCTTTTAAAAAATGGGTGATGCCCAAAATTTATGTGCCCGAAATGCTCAATGTCT 82085
QY 622 ctacacttagcctgtgatgaggagagagtgaagaagcaaaactgctggtgtcccaagga 681
Db 82084 ATTCACATAGCTTTGCAAAAAAATTTATCTGAAAATTCGATGCGAGCTGTTCGACACGGT 82025
QY 682 gcaagatttaccattgagaataaagaagaagaacacacccctgcaagtggc 731
Db 82024 GCAGACGTTAATATTATTAGCAAGAGTGGCTTCTCTCCGTTACATTGGC 81975
Search completed: August 13, 2001, 08:30:48
Job time: 2725 sec
```


299 spMetValSerLeuLeuLeuGlyArgAsnAlaAsnValAsnLeuSerAsn 315
604 ACTGAGGTAACACTCTCTACACTTACCTGTGATGAGGAGAGAGTGA 653
316 LysSerGlyLeuThrProLeuHisLeuAlaGlnGluAspArgValAs 332
654 AGAAGCAAACTGCTGTGCTCCCAAGAGCAAGTATTTACATTGAGATA 703
332 nValAlaGluValLeuValAsnGlnGlyAlaHisValAspAlaGlnThrL 349
704 AAGAAAGAAAGACACCCCTGCAAGTG 729
349 ysMetGlyThrProLeuHisVal 357
seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.us-09-172-977-4

seq_documentation_block:
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PE-0615 US
; CURRENT APPLICATION NUMBER: US/09/172, 977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-172-977-4

alignment_scores:
Quality: 294.00 Length: 232
Ratio: 2.146 Gaps: 2
Percent Similarity: 59.052 Percent Identity: 31.897
alignment_block:
US-09-509-775-1 x US-09-172-977-4 ..
Align seg 1/1 to: US-09-172-977-4 from: 1 to: 1839

136 GCCTACAGCGGAAGCTGGAAGAGTGAAGAGAGTATTCTGCGCGATAA 185
472 AlaArgAlaGlyGlnValGluValValArg...CysLeuLeuArgAsnG1 487
186 ATCCCTGGCTACTAGAACTGACGACGACAGCAACTGCTGCTGCTGGG 235
487 yAlaLeuValAspAlaArgAlaArgGluGlnThrProLeuHisIleA 504
236 CATGCTCAGTCGACATACAGAAATTTGTAATTTTGTGCAA..... 279
504 laSerArgLeuGlyLysThrGluIleValGlnLeuLeuGlnHisMet 520
279 279
521 AlaHisProAspAlaAlaThrThrAsnGlyThrProLeuHisIleSe 537
280CTGGAG 286
537 rAlaArgGluGlnValAspValAlaSerValLeuLeuGluAlaGlyA 554
287 TGCCAGTGAATATAAAGACGATGAGGTGCTCTCTTACATTGCG 336
554 laAlaHisSerLeuAlaThrLysLysGlyPheThrProLeuHisValAla 570

337 GCCTTCTGCTGGCGGATGAGATTGTAAAGCCCTTCTGGGAAAGGTGC 386
571 AlaLysThrGlySerLeuAspValAlaLysLeuLeuGlnArgAla 587
387 TCAAGTGAATGCTGTCAATAAATGGCTGTACTCCCTTACATTATGCG 436
587 aAlaAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAla 604
437 CTTCGAAAACACAGCATGATGCTGTCTATGTTACTTGAAGCGGGCT 486
604 laHisThrAspAsnGlnLysValAlaLeuLeuGlnLysGlyAla 620
487 AATCCAGATGCTAAGGACCATATTAGGCTTACACCAATGACCGGCGC 536
621 SerProHisAlaThrAlaLysAsnGlyThrProLeuHisIleAla 637
537 AGCCAAGGGTAATGAAGATGATTATCTCTCTGCTACTACAAAGCAT 586
637 aLysLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnThrGlyAla 654
587 CGCAAAACATCCAAAGACACTGAGGTAACACTCTCTACACTTAGCCTGT 636
654 luthrAsnIleValThrLysGlnGlyValThrProLeuHisLeuAlaSer 670
637 GATGAGGAGAGAGTGAAGAAAGCAAACTGCTGTGCTCCCAAGAGCAAG 686
671 GlnGluGlyHisThrAspMetValThrLeuLeuAspLysGlyAlaAs 687
687 TATTACATTGAGATAAAGAAAGACACCCCTGCAAGTGCC 732
687 nIleHisMetSerThrLysSerGlyLeuThrSerLeuHisLeuAla 702
seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.us-09-065-474-139

seq_documentation_block:
; Sequence 139, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: Protein
US-09-065-474-139

alignment_scores:
  Quality: 287.00      Length: 200
  Ratio: 2.142        Gaps: 2
  Percent Similarity: 67.000  Percent Identity: 36.500

alignment_block:
US-09-509-775-1 x US-09-065-474-139 ..
Align seg 1/1 to: US-09-065-474-139 from: 1 to: 352

154 CAAGAGTTGAAGGAGAGATTCTGGCC.....GA 182
   :|||:|||||:|||||: |||
86  LysGluGlyGlnGluValAlaAlaLeuMetAspHisGlyThrAs 102
   :|||:|||||:|||||: |||
183 TAAATCCCTGGCTACTAGAACTGACCAGGACGAGCACTGCATTGCACCT 232
   :|||:|||||:|||||: |||
102 pLysThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 116
   :|||:|||||:|||||: |||
233 GGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTCTTGCACCTT 282
   :|||:|||||:|||||: |||
116 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuLeuGluArg 132
   :|||:|||||:|||||: |||
283 GGAGTCCAGTGAATGATAAGACGATCCAGGTTGGTCTCTCTTCATAT 332
   :|||:|||||:|||||: |||
133 GlyThrProValAspLeuGluGlyLysAsnGlnValThrProLeuHisVa 149
   :|||:|||||:|||||: |||
333 TGGCGCTCTGCTGGCGGGATGAGATTGTAAAGCCCTTCTGGGAAAG 382
   :|||:|||||:|||||: |||
149 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 166
   :|||:|||||:|||||: |||
383 GTGCTCAAGTGAATGCTCAATCAAAATGCTGTACTCTTACATATAT 432
   :|||:|||||:|||||: |||
166 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisL 182
   :|||:|||||:|||||: |||
433 GCAGCTTCGAAACAGGCATGATGATGCTCATGTTGTTACTGGAAGCGG 482
   :|||:|||||:|||||: |||
183 AlaAlaLysLysAsnGlnMetAspLeuAlaSerThrLeuLeuHisTyrLy 199
   :|||:|||||:|||||: |||
483 GGCTAAATCCAGATGAGGACCATATTGAGGCTACAGCAATGCACGGG 532
   :|||:|||||:|||||: |||
199 sAlaAsnAlaAsnAlaGluSerLysAlaGlyPheThrProLeuHisLeuA 216
   :|||:|||||:|||||: |||
533 CAGCAGCCAGGGTAATGTTGAAGATGATTCATATCCTCTCTACTACAAA 582
   :|||:|||||:|||||: |||
216 LaAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuLeuGluAsnGly 232
   :|||:|||||:|||||: |||
583 GCATCCCAACATCCAGACACTGAGGGTAACACTCTCTTACTACTAGC 632
   :|||:|||||:|||||: |||
233 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeuCy 249
   :|||:|||||:|||||: |||
633 CTGTGATGAGGAGAGAGTGGGAAGCAAACTGCTGGTGTCCCAAGGAG 682
   :|||:|||||:|||||: |||
249 sAlaGlnGluAspArgValSerValAlaGluGluLeuValLysGluAsnA 266
   :|||:|||||:|||||: |||
683 CAAGTATTACATTGAGAATAAAGCAAAAGACCCCTTCCAAAGTGGCC 732
   :|||:|||||:|||||: |||
266 LaAlaLeuAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 282
   :|||:|||||:|||||: |||

seq_name: /cgn2_5/ptodata/2/1aa/5B_COMB.pep.US-09-031-485-33

seq_documentation_block:
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Biehm, E. Scott
; TITLE OF INVENTION: DIFOLIPILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
```

```
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-33

alignment_scores:
  Quality: 287.00      Length: 200
  Ratio: 2.142        Gaps: 2
  Percent Similarity: 67.000  Percent Identity: 36.500

alignment_block:
US-09-509-775-1 x US-09-031-485-33 ..
Align seg 1/1 to: US-09-031-485-33 from: 1 to: 1745

154 GAAGAGTTGAAGGAGAGATTCTGGCC.....GA 182
   :|||:|||||:|||||: |||
521 LysGluGlyGlnGluValAlaAlaLeuMetAspHisGlyThrAs 537
   :|||:|||||:|||||: |||
183 TAAATCCCTGGCTACTAGAACTGACCAGGACGAGCACTGCATTGCACCT 232
   :|||:|||||:|||||: |||
537 pLysThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 551
   :|||:|||||:|||||: |||
233 GGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTCTTGCACCTT 282
   :|||:|||||:|||||: |||
551 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuLeuGluArg 567
   :|||:|||||:|||||: |||
283 GGAGTCCAGTGAATGATAAGACGATCCAGGTTGGTCTCTCTTTCATAT 332
   :|||:|||||:|||||: |||
568 GlyThrProValAspLeuGluGlyLysAsnGlnValThrProLeuHisVa 584
   :|||:|||||:|||||: |||
333 TGGCGCTTCTGCTGGCGGGATGAGATTGTAAAGCCCTTCTGGGAAAG 382
   :|||:|||||:|||||: |||
584 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 601
   :|||:|||||:|||||: |||
383 GTGCTCAAGTGAATGCTCAATCAAAATGCTGTACTCTCTTACATATAT 432
   :|||:|||||:|||||: |||
601 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisL 617
   :|||:|||||:|||||: |||
433 GCAGCTTCGAAACAGGCATGATGATGCTCATGTTGTTACTGGAAGCGG 482
   :|||:|||||:|||||: |||
```

```
518 AlaAlaLysLysAsnGlnMetAspIleAlaSerThrLeuLeuHisTyrLy 634
483 GCCTAATCCACATGCTAAGGACCATTTATGAGGTACAGCAATGACCGGG 532
||||| :||||:||||| :||| :|||||
634 sAlaAsnAlaAsnAlaGluSerLysAlaGlyPheThrProLeuHisLeuA 651
||||| :||||:||||| :||||| :|||||
533 CAGCAGCAAGGGTAACTTGAAGATGATTCATATCCTCTCTACTACAAA 582
||||| :||||:||||| :||||| :|||||
651 laAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuIleGluAsnGly 667
||||| :||||:||||| :||||| :|||||
583 GCATCCCAAAATCCCAAGACACTGAGGGTAACTCTCTACTACTTACG 632
||||| :||||:||||| :||||| :|||||
668 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeuCy 684
||||| :||||:||||| :||||| :|||||
633 CTGTCTAGCAGAGAGGGAAGAGCAAACTGCTGTGTGCCAAGGAG 682
||||| :||||:||||| :||||| :|||||
684 saLaGlnGluAspArgValSerValAlaGluGluLeuValLysGluAsnA 701
||||| :||||:||||| :||||| :|||||
683 CAAGTATTATTCATTCAGATAAAGAGAAAGACACCCCTGCAAGTGGCC 732
||||| :||||:||||| :||||| :|||||
701 laAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 717
||||| :||||:||||| :||||| :|||||
seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.us-08-847-429A-33

seq_documentation_block:
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-33

alignment_scores:
Quality: 287.00 Length: 200
Ratio: 2.142 Gaps: 2
Percent Similarity: 67.000 Percent Identity: 36.500

alignment_block:
```

```
US-09-509-775-1 x US-08-847-429A-33
Align seg 1/1 to: US-08-847-429A-33 from: 1 to: 1745
154 GAAGAGTTGAAGGAGAGTATTCCTGGCC.....GA 182
||||| :||||:||||| :||| :|||||
521 LysGluGlyGlnGluGluValAlaAlaLeuMetAspHisGlyThrAs 537
||||| :||||:||||| :||||| :|||||
183 TAAATCCCTGCTACTAGAACTGACCAGGACAGCAAACTGCAATTCAC 232
||||| :||||:||||| :||||| :|||||
537 PLYSThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 551
||||| :||||:||||| :||||| :|||||
233 GGCATGCTCAGTGGACATACAGAAATTTGAAATTTTGTTCGCACT 282
||||| :||||:||||| :||||| :|||||
551 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuLeuGlu 567
||||| :||||:||||| :||||| :|||||
283 GGAGTCCCACTGAATGATAAGACGATGCGAGTTGGTCTCTCTCAT 332
||||| :||||:||||| :||||| :|||||
568 GlyThrProValAspIleGluGlyLysAsnGlnValThrProLeuHis 584
||||| :||||:||||| :||||| :|||||
333 TGGCGCTCTGCTGGCGGATGAGATTGTAAGGCCCTTCTGGGAAAG 382
||||| :||||:||||| :||||| :|||||
584 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 601
||||| :||||:||||| :||||| :|||||
383 GTGCTCAAGTGAATGCTCAATCAAAATGCTGTACTCTCTTACATTAT 432
||||| :||||:||||| :||||| :|||||
601 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisIle 617
||||| :||||:||||| :||||| :|||||
433 GCAGCTTCGAAACACGCGATGCGTGTCTGCTACTGGAAGCGG 482
||||| :||||:||||| :||||| :|||||
618 AlaAlaLysLysAsnGlnMetAspIleAlaSerThrLeuLeuHisTyr 634
||||| :||||:||||| :||||| :|||||
483 GGCTAATCCAGATGCTAAGGACCAATATGAGGCTACAGCAATGCA 532
||||| :||||:||||| :||||| :|||||
634 sAlaAsnAlaAsnAlaGluSerLysAlaGlyPheThrProLeuHis 651
||||| :||||:||||| :||||| :|||||
533 CAGCAGCAAGGGTAACTTGAAGATGATTCATATCCTCTCTACTACA 582
||||| :||||:||||| :||||| :|||||
651 laAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuIleGluAsn 667
||||| :||||:||||| :||||| :|||||
583 GCATCCCAAAATCCCAAGACACTGAGGGTAACTCTCTACTACTTAC 632
||||| :||||:||||| :||||| :|||||
668 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeu 684
||||| :||||:||||| :||||| :|||||
633 CTGTGATGAGGAGAGTGGGAAGAGCAAACTGCTGTGTCCCAAGGAG 682
||||| :||||:||||| :||||| :|||||
684 saLaGlnGluAspArgValSerValAlaGluGluLeuValLysGlu 701
||||| :||||:||||| :||||| :|||||
683 CAAGTATTATTCATTCAGATAAAGAGAAAGACACCCCTGCAAGTGG 732
||||| :||||:||||| :||||| :|||||
701 laAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisVal 717
||||| :||||:||||| :||||| :|||||
seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.us-09-065-474-33

seq_documentation_block:
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
```

[illegible]

```
1 .....|
552 aAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAlaA 569
437 CTTGCAAAACAGCATGAGTCCCTGCTCATCTGTTACTGGAGCGGGCT 486
569 laHisTyAspAsnGlnLysValAlaLeuLeuLeuGluLysGlyAla 585
487 AATCAGATGCTAAGGACCAATATGAGCTACAGCAATGCACCGGGCAGC 536
586 SerProHisAlaThrAlaLysAsnGlyTyThrProLeuHisIleAlaAl 602
537 AGCCAAGGGTAACCTGAGATGATCATATCTCTCTACTACTACAAAGCAT 586
602 aLysLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnTyrglyAlaG 619
587 CCACAAACATCCAGACACTGAGGGTACACTCTCTACACTTACCTGTGT 636
619 luThrAsnThrValThrLysGlnGlyValThrProLeuHisLeuAlaSer 635
637 GATGAGGAGAGTGGGAAGCAAGCAAACTGCTGTGTCCTCCCAAGGCAAG 686
636 GlnGluGlyHisThrAspMetValThrLeuValLeuGluLysGlyAlaAs 652
687 TATTTCATTCAGATAAAGAAAGACACACCCTGCAAGTGGCC 732
652 nileHisMetSerThrLysSerGlyLeuThrSerLeuHisLeuAla 667
seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-031-485-38

seq_documentation_block:
; Sequence 38, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESS: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-38
```

```
alignment_scores:
  Quality: 271.00      Length: 189
  Ratio: 2.134        Gaps: 0
  Percent Similarity: 67.196      Percent Identity: 35.450

alignment_block:
US-09-509-775-1 x US-09-031-485-38 ..
Align seg 1/1 to: US-09-031-485-38 from: 1 to: 302

166 GAGAGTATTCTGGCCGATAAATCCCTGGCTACTAGAACTGCACGACGAC 215
||| .....|
57 LeuLeuLeuLysTyHisAlaAlaIleGluAlaThrThrGluSerG1 73

216 CAGAACTGCATTGCACCTGGGCATGCTCAGCTGCACATACAGAAATTTG 265
: : : : :|
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90

266 AATTTTGTTCACACTTGGAGTCCAGTGAATGATAAAGACGATCGAGGT 315
||| .....|
90 leTyLeuLeuGlnGlnGlyAlaAsnAlaAsnValAlaThrValArgGly 106

316 TGGTCTCCTCTTCATATGCGGCTTCTGTCGGCGGATCAGATTGTAA 365
: : : : :|
107 GluThrProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValAr 123

366 AGCCCTCTCGGAAAGGTCTCAAGTGAATGCTCAATCAATAAATGGCT 415
: : : : :|
123 gValLeuValArgAsnGlyAlaGlnValAlaAlaAlaArgGluLeuG 140

416 GTACTCCTTACATTTAGCTTCCAGTTCGAAACACAGCATGAGTCGTGC 465
||| .....|
140 lnThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIle 156

466 ATGTTACTGGAAGCGGGCTAATCCAGATGCTTAAGACCATATGAGGC 515
: : : : :|
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeuTy 173

516 TACAGCAATGCACCGGCAGCAGCAAGGTAACTTGAAGATGATTCATA 565
||| .....|
173 rThrLeuLeuHisIleAlaAlaLysGluGlnGluValAlaAlaAl 190

566 TCCTTCTGTACTACAAAGCATCCACAAACATCCAAAGACACAGGTGTAAC 615
||| .....|
190 leLeuIleAspHisGlySerAspLysThrLeuLeuThrLysGlyPhe 206

616 ACTCCTCTACACTTAGCTGTGTGATGAGGAGAGTGGGAAGACAAACT 665
||| .....|
207 ThrProLeuHisLeuAlaAlaLysTyThrGlyAsnLeuProValAlaLysLe 223

666 GCTGGTGTCCCAAGGAGCAAGTATTTACATTGAGATAAAGAAAGAAAGA 715
||| .....|
223 uLeuLeuGluArgGlyThrLeuValAlaAspIleGluGlyLysAsnGlnValT 240

716 CACCCCTGCAAGTGGCC 732
||| .....|
240 hrProLeuHisValAla 245

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-847-429A-38

seq_documentation_block:
; Sequence 38, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-847-429A-38

```

alignment_scores:		
Quality:	271.00	Length: 189
Ratio:	2.134	Gaps: 0
Percent Similarity:	67.196	Percent Identity: 35.450

```
alignment_block:
US-09-509-775-1 x US-08-847-429A-38 ..
```

Align seg 1/1 to: US-08-847-429A-38 from: 1 to: 302

166 GAGAGTATTCTGGCCGATAAATCCTCGGTACTAGAACTGACCAGGACAG 215
||| ||||| : : : ||| : : : : :
57 GluLeuLeuLysThrHisAlaAlaIleGlAlaAlaThrThrGluSerGI 73
:
216 CAGAACTGCATTGCTACTGGCATCTCAGCTGGACATACAGAATAATTGTTC 265
: : : ||||||| ||||| : : : ||| : : : |||||
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90
:
266 AATTTTGTGCAACTTCGAGTCCGACTGAATGATTAAGACGATGCAGGT 315
: : : ||||| : : : : : ||| : : : |||
90 LeTyrLeuLeuGlnGlnGlyAlaAsnAlaAsnValAlaThrIvalArgGly 106
:
316 TGGTCTCTCTTCATATTTCGGCTCTCTCTGGCGGGATCAGATTCATAA 365
: : : ||||| : : : ||||| : : : ||||| : : : |||||
107 GluThrProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValAr 123
:
366 AGCCCTTCTGGAAAAGTGCTCAAGTCAATGTGTCTCAATCAAANAATGGCT 415
: : : ||||| : : : ||||| : : : ||||| : : : : :
123 gValLeuValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuG 140
:
416 GTACTCCCTTACATTATCCAGCTTCGAAAAACAGGCATGAGATCGCTGTC 465
||||| ||||| : : : : : : : : : : : : : : :
140 InThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIle 156
:
466 ATGTCTACTGGAAGGGCGGCTTAATCCAGATGCTAAGGCCATTATTCAGGC 515
: : : ||||| : : : ||||| : : : ||| : : : : :
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeuTy 173
:
516 TCACAGCATGCAACGGGCGAGCGAACAGGTTAACTTGAAGATGATTCATA 565
||| ||||| ||||| : : : ||||| : : : ||||| : : : |||||
173 rThrLeuLeuHisIleAlaLalYsGluGlyGlnGluGluValAlaAlar 190

alignment_scores:		
Quality:	271.00	Length: 189
Ratio:	2.134	Gaps: 0
Percent Similarity:	67.196	Percent Identity: 35.450

```
alignment_block:
US-09-509-775-1 x US-09-065-474-38
```

Align seq 1/1 to: US-09-065-474-38 from: 1 to: 302

166 GAGAGTATTCTGGCCGATAAATCCCTGGCTACTAGAACTGACCAGGACAG 215

```

||| .....|||
57 GluLeuLeuLeuLysTyrHisAlaAlaIleGluAlaThrThrGluSerGI 73
216 CAGAACTGCATTCAGCTGGGATGCTCAGCTGGACATACAGAAATTTGG 265
: : : ||||| ||||| : : : |||||
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90
266 AATTTTGTTCGAACCTGGAGTGGCCAGTGAATGATAAGACGATGCAGGT 315
: : : ||||| ||||| : : : |||||
90 LeuTyrLeuGluGlnGlyAlaAsnAlaAsnValAlaThrValArgGly 106
316 TGGTCTCTCTTCATATTGGCTTCTGCTGGCGGGATGAGATTGTAAA 365
: : : ||||| ||||| : : : |||||
107 GluThrProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValar 123
366 AGCCCTCTCTGGGAAAGTGTCTCAAGTGAATGCTCAATCAAAATGGCT 415
: : : ||||| : : : ||||| : : : |||||
123 gValLeuValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuG 140
416 GTACTCCCTTACATTATGCAGCTTCGAAACACGACATGAGATCGCTGC 465
: : : ||||| ||||| : : : |||||
140 InThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIle 156
466 ATGTACTGGAAGCGGGCTTAATCCAGATGCTAGGACCATTTATGAGGC 515
: : : ||||| : : : ||||| : : : |||||
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeuty 173
516 TACAGCATGCACCGGCGAGCGAGCAAGGTAACTTGAAGATGATTCATA 565
||| : : : ||||| : : : ||||| : : : |||||
173 rThrLeuLeuHisIleAlaAlaLysGluGlyGlnGluValAlaAlaI 190
566 TCCTCTGTCTACAAAGCATCCCAACATCCAAAGACACTGAGGTAAC 615
||||| : : : ||||| : : : |||||
190 leuIleAspHisGlySerAspLysThrLeuLeuThrLysLysGlyPhe 206
616 ACTCCTCTACACTTACCTGTGATGAGGAGAGTGAAGGAAAGAAAAC 665
||||| ||||| ||||| ||||| : : : |||||
207 ThrProLeuHisLeuAlaAlaLysTyrGlyAsnLeuProValAlaLysLe 223
666 GCTGGTGTCCCAAGGACAAGTATTACATTCAGATAAAGAAAGAAAGA 715
223 uLeuLeuGluArgGlyThrLeuValAspIleGluGlyLysAsnGlnVal 240
716 CACCCCTGCAAGTGGCC 732
240 hrProLeuHisValAla 245

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep.US-09-031-485-23
seq_documentation_block:
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIOFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-23

alignment_scores:
Quality: 262.00 Length: 218
Ratio: 1.912 Gaps: 4
Percent Similarity: 62.844 Percent Identity: 33.945

alignment_block:
US-09-509-775-1 x US-09-031-485-23 ..
Align seg 1/1 to: US-09-031-485-23 from: 1 to: 303

136 GCCTACAGCGGGAAGCTGGAAGAGTTGAAGAGAGAGTATTCTGCCCGATAA 185
||| : : : ||||| : : : |||||
16 AlaHisCysGlyHisValArgValAlaLys.....LeuLeuLeuAspAr 30
186 ATCCCTGGCTACTAGAAGTACCAGGACAGCAGA.....A 220
: : : ||||| : : : |||||
30 GASD.....AlaAspProAsnAlaArgAlaLeuAsnGlyPheT 43
221 CTGCATTGCACTGGGCATGCTCAGCTGGACATACAGAAATTTGGAATTT 270
||| ||||| ||||| : : : |||||
43 hrProLeuHisIleAlaCysLysLysAsnArgIleLysIleValGluLeu 59
271 TTGTTGCAACTTGAGTCCCGCAGTGAATGATAAGACATGACGAGTTGGTC 320
||||| : : : ||||| : : : |||||
60 LeuLeuLysTyrHisAlaAlaIleGluAlaThrThrGluSerGlyLeuSe 76
321 TCCTCTTCATATTGGGCTTCTGCTGGCGGAGTGAATTTGTAAGAGCC 370
||||| ||||| ||||| |||||
76 rProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValIleTyrL 93
371 TTCTGGGAAAAAGGTGCTCAAGTGAATGCTGTCATCAAAATGGCTGACT 420
||||| : : : ||||| : : : |||||
93 euLeuGlnGlnGlyAlaAsnAlaAspValAlaThrValArgGlyGluThr 109
421 CCCTTACATTTATGAGCTTCCAAAACAGGATGAGATCGCTGTCATGTT 470
||||| ||||| ||||| : : : |||||
110 ProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValArgValle 126
471 ACTGGAAGCGGGGCTAATCCAGATGCTTAAGGACCATTTATGAGGCTACAG 520
||| : : : ||||| : : : |||||
126 uValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuGlnThrP 143
521 CAATGCACCGGCGAGCAGCCCAAGGTAACCTTGAAGATGATTCATATCCTT 570
: : : ||||| ||||| : : : |||||
143 roLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIleLeuLeu 159
571 CTGTACTACAAGCATCCACAACATCCAGACACTGAGGGTTAACACTCC 620
||| : : : ||||| |||||
160 LeuGlnAlaAsnAlaSerProAsnAlaAlaThrArgAspLeuTyrThrPr 176
621 TCTACACTTAGCCTGTGATGAGGAGAGTGGAGAAGCAAAACACTGCTGG 670

```



```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-23

alignment_scores:
    Quality: 262.00      Length: 218
    Ratio: 1.912        Gaps: 4
    Percent Similarity: 62.844      Percent Identity: 33.945

alignment_block:
US-09-509-775-1 x US-09-065-474-23  ..
Align seg 1/1 to: US-09-065-474-23 from: 1 to: 303

136 GCCTACAGCGGAACTGGAAGAGTTGAAGGAGATATTCTGGCGGATAA 185
||||| ||| ||||| ||| ||||| ||||| ||||| |||||
16 AlaHisCysGlyHisValArgValAlaLys.....LeuLeuAspAr 30
||||| ||||| ||||| ||||| ||||| ||||| |||||
186 ATCCCTGGCTACTAGAACTGACCAAGGACGACGAGA.....A 220
||||| ||||| ||||| ||||| ||||| ||||| |||||
30 gAsn.....AlaAspProAsnAlaArgAlaLeuAsnGlyPheT 43
||||| ||||| ||||| ||||| ||||| ||||| |||||
221 CTGCATCTGACTGGGATGCTCAGCTGACGATACACAGAAATTTGAATTT 270
||||| ||||| ||||| ||||| ||||| ||||| |||||
43 hrProLeuHisIleAlaCysLysLysAsnArgIleLysIleValGluLeu 59
||||| ||||| ||||| ||||| ||||| ||||| |||||
271 TTGTGCAACTTGGAGTCCAGTGAATGATATAAGGATGACAGTTGCTC 320
||||| ||||| ||||| ||||| ||||| ||||| |||||
60 LeuLeuLysTyrHisAlaIleGluAlaThrThrGluSerGlyLeuSe 76
||||| ||||| ||||| ||||| ||||| ||||| |||||
321 TCCTCTTCATATTCGGCTTCTGCTGGCGGGATGAGATTGTAAGACCC 370
||||| ||||| ||||| ||||| ||||| ||||| |||||
76 rProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValIleTyrL 93
||||| ||||| ||||| ||||| ||||| ||||| |||||
371 TTCGGGAAAGTCTCAACTGAATGCTGTCAATCAAATGGCTGACT 420
||||| ||||| ||||| ||||| ||||| ||||| |||||
93 euLeuGlnGlnGlyAlaAsnAlaAspValAlaThrValArgGlyGluThr 109
||||| ||||| ||||| ||||| ||||| ||||| |||||
421 CCCTTACATATTCAGCTCTCAAAAACAGCATGAGATCGCTCATGTT 470
||||| ||||| ||||| ||||| ||||| ||||| |||||
110 ProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValArgValIle 126
||||| ||||| ||||| ||||| ||||| ||||| |||||
471 ACTGGAAGCGGGCTAATCCAGATGCTAAGGACCATATGAGGCTACAG 520
||||| ||||| ||||| ||||| ||||| ||||| |||||
126 uValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuGlnThrP 143
||||| ||||| ||||| ||||| ||||| ||||| |||||
521 CAATGCAACGGGACGACCAAGGTAACCTTGAAGATGATTTCATATCCCT 570
||||| ||||| ||||| ||||| ||||| ||||| |||||
143 roLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIleLeuLeu 159
||||| ||||| ||||| ||||| ||||| ||||| |||||
571 CTGTACTACAAAGCATCCCAAAACATCCAGACACTGAGGGTAACACTCC 620
||||| ||||| ||||| ||||| ||||| ||||| |||||

```

THIS PAGE BLANK (USPTO)

OM of: US-09-509-775-1 to: A_Geneseq_0601.* out_format : pfs

Date: Aug 13, 2001 8:30 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framet_n2p_model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09509775/runat_13082001_074527_13026/app_query.fasta_1.848
-DB=A_Geneseq_0601 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000
-XGAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09509775_@CGNL_1_51 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1
```

Search information block:

```
Query: US-09-509-775-1
Query length: 780
Database: A_Geneseq_0601.*
Database sequences: 412676
Database length: 60623988
Search time (sec): 74.500000
```

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:AAW15483	226	1164.00	226	8.3e-119	226
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02430	226	1164.00	226	8.3e-119	226
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02432	231	1110.00	216.77	6.9e-113	231
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02431	231	1105.00	215.94	2.4e-112	231
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW02431	1166	298.50	547.76	1.1e-23	1166
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66287	522	295.50	550.95	1.6e-23	522
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66278	1169	295.50	541.83	2.3e-23	1169
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66288	1169	295.50	541.83	2.3e-23	1169
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66290	1262	295.50	540.97	2.4e-23	1262
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66294	1385	295.50	539.91	2.5e-23	1385
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66295	1166	295.50	537.93	3.8e-23	1166
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66295	991	291.50	535.83	5.9e-23	991
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW7023	1166	290.50	532.03	8.1e-23	1166
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW72589	1166	287.50	526.13	1.7e-22	1166
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11616	352	287.00	538.69	1.1e-22	352
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW0608	1745	287.00	520.59	2.4e-22	1745
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW76776	1745	287.00	520.59	2.4e-22	1745
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11589	1745	287.00	520.59	2.4e-22	1745
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44403	673	281.50	520.54	6.2e-22	673
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44404	949	281.50	516.66	7.2e-22	949
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW27212	1327	281.50	512.87	8.4e-22	1327
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44402	1327	281.50	512.87	8.4e-22	1327
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66279	1327	281.50	512.87	8.4e-22	1327
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66297	1181	281.00	513.20	9.0e-22	1181
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW70609	302	271.00	508.96	6.1e-21	302
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW6777	302	271.00	508.96	6.1e-21	302
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11590	302	271.00	508.96	6.1e-21	302
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW53572	1030	270.50	494.10	1.2e-20	1030
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW41378	976	269.50	492.75	1.5e-20	976
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW53571	976	269.50	492.75	1.5e-20	976
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW41377	1074	269.00	490.68	1.8e-20	1074
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW05734	756	265.00	486.79	4.2e-20	756
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66286	784	265.00	486.37	4.2e-20	784
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66285	907	265.00	484.73	4.5e-20	907
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW48574	303	262.00	491.22	5.9e-20	303
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW70606	303	262.00	491.22	5.9e-20	303
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW6774	303	262.00	491.22	5.9e-20	303
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11587	978	260.50	475.03	1.5e-19	978
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW42288	763	259.50	475.87	1.7e-19	763

```
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW69163 + 259.50 475.53 1.7e-19 786
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW76079 + 259.50 475.52 1.7e-19 787
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW56018 + 259.50 475.52 1.7e-19 787
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW12893 + 252.50 467.92 7.8e-19 456
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW27402 + 252.50 467.92 7.8e-19 456
```

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:AAW15483

seq_documentation_block:

ID AAW15483 standard; Protein: 226 AA.

XX AAW15483;

DT 17-JUN-1997 (first entry)

XX Human P28.

XX Human; proteasome; P28; diagnosis; malignant tumour.

XX Homo sapiens.

XX JP09075085-A.

PD 25-MAR-1997.

PF 13-SEP-1995; 95JP-0235052.

PR 13-SEP-1995; 95JP-0235052.

PA (SAGA) SAGAMI CHEM RES CENTRE.

XX WPI; 1997-239267/22.

DR N-PSDB; AAT66424-25.

XX Human 26S proteasome constituting component protein - useful in the diagnosis of e.g. malignant tumour

PS Claim 1; Page 6-7; 9pp; Japanese.

XX This sequence represents the human proteasome component protein P28. The protein, P28, is useful for the diagnosis and treatment of various diseases caused by proteasomes such as malignant tumour.

SQ Sequence 226 AA;

alignment_scores:

Quality: 1164.00 Length: 226

Ratio: 5.150 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-509-775-1 x AAW15483

Align seg 1/1 to: AAW15483 from: 1 to: 226

97 ATGAGGGGTGCTGTCTACCTAAATGGTCTGCAACCTGGCCTACACGG 146

||||| 1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG1 17

||||| 147 GAAGCTGGAGAGCTTCAAGGAGAGTATCTGGCCGATAAATCCCTGGCTA 196

||||| 17 yLysLeuGluGluLeuLysSerIleLeuAlaAspLysSerLeuAla 34

||||| 197 CTAGAAGTCAACAGGACAGCAGCAACCTGGCCTACCTGGCCTACAGCT 246

||||| 34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

||||| 247 GGACATACAGAAATCTGTAATTTTGTGGCAACTGGAGTGGCAGTGAA 296

||||| 51 GlyHisThrGluIleValGluPheLeuLeuGluGlyValProValAs 67

||||| 297 TGATAAGACGATGCGAGTGGTCTCTCTTCATATTGCGGCTTCTGCTG 346

201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217

747 TTTAATACTCAAGAGAAATGGTGGAGGT 774

217 yLeuIleLeuLysArgMetValGluGly 226

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.AAY02432

seq_documentation_block:

ID AAY02432 standard; Protein; 231 AA.

XX AC AAY02432;

DT 14-JUL-1999 (first entry)

XX Rat gankyrin protein.

DE Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.

XX Rattus sp.

XX WO9918201-A1.

PD 15-APR-1999.

XX 02-OCT-1998; 98WO-JP04467.

PR 03-OCT-1997; 97JP-0286214.

XX (FUJII/) FUJITA.

XX Fujita J;

DR WPI; 1999-277266/23.

DR N-PSDB; AAX35854.

XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism

PS Claim 11; Page 76-78; 11pp; Japanese.

CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents rat gankyrin.

XX Sequence 231 AA;

alignment_scores:

Quality: 1110.00 Length: 225

Ratio: 4.978 Gaps: 0

Percent Similarity: 99.111 Percent Identity: 95.111

alignment_block:

US-09-509-775-1 x AAY02432 ..

Align seg 1/1 to: AAY02432 from: 1 to: 231

97 ATGAGGGGTGTGTCTACCTTAATGGTGTGCAACCTGGCTACACGG 146

1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaIleAsnG1 17

147 GAAGCTGGAAGAGTTGAAGGAGAGTATTCTGGCGGATAAATCCCTGGCTA 196

17 yLysLeuAspGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34

197 CTAGAAGTACACGAGCAGCAGACTGCATTGGCATTGGCGATGCTCAGCT 246

34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

247 GCACATACAGAAATTTGTAATTTTGTGCAACTTGGAGTGCAGTGAA 296

51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67

297 TGATAAAGACGATCGAGTTGGTCTCTCTTCATATTTGCGGCTTCTGCTG 346

67 nGluLysAspAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84

347 GCCGGGATGAGATTGTAAGCCCTTCTGGGAAAAGGTGCTCAAGTGAAT 396

84 lyArgAspGluIleValLysAlaLeuLeuIleLysGlyAlaGlnValasn 100

397 GCTGTCATCAATCAAAATGCTGTACTCTCTTACATTATGCAGCTTCGAAAA 446

101 AlaValasnGlnasnGlyCysThrAlaLeuHisIleAlaAlaSerLysas 117

447 CAGGCATGAGATCGCTGTCATGTTTACTGGAAAGCGGGGCTTAATCCAGATG 496

117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAsnProAspA 134

497 CTAAGGACCATTTATGAGGCTACAGCAATGCACCGGCGAGCCCAAGGCT 546

134 laLysAsnHisTyrAspAlaThrAlaMetHisArgAlaAlaAlaLysGly 150

547 AACTTGAAGATGATTTCATATCTCTTCTACTACAAAGCATCCACAACAT 596

151 AsnLeuLysMetValHisIleLeuLeuPheTyrLysAlaSerThrAsnI1 167

597 CCAAGACACTGAGGTAACACTCTCTACACTTACCTGTGATGAGGAGA 646

167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysaspGluGluA 184

647 GAGTGGGAAGCAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTTTACATT 696

184 rgValGluGluAlaLysLeuLeuValThrGlnGlyAlaSerIleTyrIle 200

697 GAGATAAAGAAAGAAAGACACACCTGCAAGTGGCCAAAGGTGGCTGGG 746

201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217

747 TTTAATACTCAAGAGAAATGGTGGAA 771

217 yLeuIleLeuLysArgIleValGlu 225

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.AAY02431

seq_documentation_block:

ID AAY02431 standard; Protein; 231 AA.

XX AC AAY02431;

DT 14-JUL-1999 (first entry)

DE Mouse gankyrin protein.

XX Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.

XX Mus sp.

XX WO9918201-A1.

PD 15-APR-1999.

XX 02-OCT-1998; 98WO-JP04467.

PR 03-OCT-1997; 97JP-0286214.

XX (FUJII/) FUJITA.

XX Fujita J;

XX WPI: 1999-277266/23.
DR N-PSDB; AAX35853.
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 6; Page 73-74; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents mouse gankyrin.
XX
SQ Sequence 231 AA;

alignment_scores:
Quality: 1105.00 Length: 225
Ratio: 4.955 Gaps: 0
Percent Similarity: 99.111 Percent Identity: 93.333
alignment_block:
US-09-509-775-1 x AAY02431 ..
Align seg 1/1 to: AAY02431 from: 1 to: 231

97 ATGAGGGGTGTGTTCTTAACCTAATGTCACCTGGCTACAGCGG 146
|||||
1 MetGluGlyCysValSerAsnMetileCysAsnLeuAlaTyr-SerG1 17
147 GAGCTGGAGAGTGAAGGAGAGTATCTGCGCCGATATCCCTGGCTA 196
|||||
17 yLysLeuAspGluLeuLysGluArgileLeuAlaAspLysSerLeuAla 34
197 CTAGAAGTACAGGAGACAGACAGAACTGCTGCTGCGGCTGCTGAGT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50
247 GGACATACAGAAATGCTGAATTTTGTGCACTTGGAGTGGCAGTGA 296
|||||
51 GlyHisThrGluLeuValGluPheLeuLeuGluLeuGluValProValas 67
297 TGATAAAGACGATGCGAGTGTGCTCTCTTCATATTGCGGCTTCTGCTG 346
|||||
67 nAspLysAspAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84
347 GCGGGGATGAGATTGTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAAT 396
|||||
84 lyArgAspGluIleValLysAlaLeuLeuValLysGlyAlaHisValasn 100
397 GCTGTCATCAAAATGCGTCTACTCCCTTACATTTATGACGCTCGAA 446
:|||||
101 SerValasnGlnAsnGlyCysThrProLeuHisTyrAlaAlaSerLysas 117
447 CAGGCATGATGCGTCTCATGTTACTGGAAGCGGGGCTAATCCAGATG 496
|||||
117 narGHisGluIleSerValMetLeuLeuGluGlyGlyAlaasnProAspA 134
497 CTAAGGACCATTTATAGGCTACAGCAATCACCGGCGAGCAGCGAGGT 546
|||||
134 lalysAspHisTyrAspAlaThrAlaMetHisArgAlaAlaLysGly 150
547 AACTTCAAGATGATTATATCTCTCTGCTACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetValHisIleLeuLeuPheTyrLysAlaSerThrAsnI1 167
597 CCAAGACACTGAGGGTAACACTCCTTACACTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluGluA 184

647 GAGTGGAAAGCAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysPheLeuValThrGlnGlyAlaSerIleTyrIle 200
697 GAGATAAAGAAAGAACACACCCCTGCAAGTGGCCAAAGTGGCCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217
747 TTTAATACTCAAGAGATGGTGA 771
|||||
217 yLeuIleLeuLysArgLeuAlaGlu 225
seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AA47022
seq_documentation_block:
ID AAB47022 standard; Protein; 1166 AA.
XX
AC AAB47022;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 173..209 /note= "Encoded by nucleotides 768..878"
FT FT Misc-difference 327..362 /note= "Encoded by nucleotides 1230..1337"
FT FT Misc-difference 375..398 /note= "Encoded by nucleotides 1374..1445"
FT FT Misc-difference 482..524 /note= "Encoded by nucleotides 1695..1823"
FT FT Misc-difference 641..677 /note= "Encoded by nucleotides 2172..2282"
XX
PN WO200077225-A1.
XX
PD 21-DEC-2000.
XX
PF 09-JUN-2000; 2000WO-US15926.
XX
PR 11-JUN-1999; 99US-0138957.
XX
PA (WHEH) WHITEHEAD INST BIOMEDICAL RES.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Chi N, Lodish HF;
XX
DR WPI: 2001-091404/10.
DR N-PSDB; AAC85294.
XX
PT New insulin signalling protein SPANK, useful for reducing body mass,
PT glucose intolerance or insulin resistance and for preventing or
PT treating obesity-related and muscle-related diseases
XX
PS Claim 3; Fig 3; 65pp; English.
XX
CC This sequence represents human SPANK. The SPANK protein comprises
CC 3 domains:
CC (a) a SAM (sterile alpha motif) domain;
CC (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic


```
505 .....CATTATGAGGCTA 517
165 pLeuAlaaspProSerAlaLysAlaValLeuThrGlyGluTyrLysLysA 182
518 CAGCAATGCACCGGAGCAGCCAGGTAAC...TTGAAGATGATTCAAT 564
182 spGluLeuLeuGluSerAlaArgSerGlyAsnGluGluLysMetMetAla 198
565 ATCTCTCTGTACTACAAGCATCCACAACATCCAAACACTGAGGGTAA 614
199 LeuLeuThrProLeuAsnValAsnCysHisAlaSerAspGlyArgLysSe 215
615 CACTCTCTACACTTAGCTGTGTAGGAGAGAGAGTGAAGAAGCAAAAC 664
215 rThrProLeuHisLeuAlaAlaGlyTyrAsnArgValLysIleValGlnL 232
665 TGCTGGTGTCCCAAGGAGCAAGTATTACATTGAGATTAAGAAGAAAG 714
232 euLeuLeuGlnHisGlyAlaAspValHisAlaLysAspLysGlyAspLeu 248
715 ACACCCCTGCAAGTGGCC 732
249 ValProLeuHisAsnAla 254
```

seq_name: /SIDSI/gcdata/geneseq/geneseq/AA2001.DAT:AA66290

seq_documentation_block:

ID: AAB66290 standard; Protein; 1262 AA.

AC AAB66290;

05-APR-2001 (first entry)

Human tankyrase2 clone consensus protein SEQ ID NO: 107.

Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging; inflammatory disorder.

Homo sapiens.

WO200100849-A1.

04-JAN-2001.

28-JUN-2000; 2000WO-US17827.

29-JUN-1999; 99US-0141582.

(ICOS-) ICOS CORP.

Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

WPI; 2001-102896/11.

N-PSDB; AAF63930.

New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

Example 2; Page 173-176; 242pp; English.

The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has CC polyADP-ribosylation activity and is involved in the modification of CC TRF1, which is a telomere-specific binding protein. The regulation of CC telomere length, in which TRF1 has a role, is linked to aging and CC cancer. The sequences are useful in the treatment of cancers and CC inflammatory disorders.

Sequence 1262 AA;

```
alignment_scores:
  Quality: 295.50      Length: 256
  Ratio: 1.824        Gaps: 4
  Percent Similarity: 63.281  Percent Identity: 31.641
alignment_block:
US-09-509-775-1 x AAB66290
..
Align seg 1/1 to: AAB66290 from: 1 to: 1262
31 CGTAGCCGGAGCGCGGCGGCTGCTCGCGCGCAAGTAG 80
111 111 111 111 111 111 111 111 111 111
95 ArgIleMetSerGlyArgCysAlaGlyGlyAlaAlaCysAlaSe 111
81 TTGCTGGGACAGCGAATGAGGGGTGTGTAACTAATG...GTCT 127
111 111 111 111 111 111 111 111 111 111
111 rAlaAlaAlaGluAlaValGluProAlaAlaArgGluPheGluAlaC 128
128 GCAACCTGGCTACAGCGGGAAGCTGGAAGAGTTGAAGAGAGTATTCTG 177
111 111 111 111 111 111 111 111 111 111
128 ysarg.....AsnGlyAspValGluArgValLysArgLeuValThr 141
178 GCCGATAAATCCCTGGCTACTAGAACTGACCAGCAGCAGCAACTGCATT 227
111 111 111 111 111 111 111 111 111 111
142 ProGluLysValAsnSerArgAspThrAlaGlyArgLysSerThrProLe 158
228 GCACCTGGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTGTTC 277
111 111 111 111 111 111 111 111 111 111
158 uHisPheAlaAlaGlyPheGlyArgLysAspValValGluTyrLeuLeu 175
278 AACTTGAGTGCCAGTGAATGAATAAGACGATGAGTGTGTCTCTCTCT 327
111 111 111 111 111 111 111 111 111 111
175 lnAsnGlyAlaAsnValGlnAlaArgAspAspGlyGlyLeuLeuProLeu 191
328 CATATTGCGCTTCTGCTGCCGGGATGAGATTGATAAAGCCCTTCTGGG 377
111 111 111 111 111 111 111 111 111 111
192 HisAsnAlaCysSerPheGlyHisAlaGluValValAsnLeuLeuLeu 208
378 AAAGGTGCTCAAGTGAATGCTCAATCAAAATGGCTGTACTCCCTTAC 427
111 111 111 111 111 111 111 111 111 111
208 gHisGlyAlaAspProAsnAlaArgAspAsnTrpAsnTrpThrProLeuH 225
428 ATTATGAGCTTCGAAACAGGAGCATGAGTGTCTGCTACTGTTACTGAA 477
111 111 111 111 111 111 111 111 111 111
225 iGluAlaAlaIleLysGlyLysIleAspValCysIleValLeuLeuGln 241
478 GCGGGGCTAATCCAGATGCTAAGGAC..... 504
111 111 111 111 111 111 111 111 111 111
242 HisGlyAlaGluProThrIleArgAsnThrAspGlyArgThrAlaLeuAs 258
505 .....CATTATGAGGCTA 517
258 pLeuAlaAspProSerAlaLysAlaValLeuThrGlyGluTyrLysLysA 275
518 CAGCAATGCACCGGAGCAGCCAGGTAAC...TTGAAGATGATTCAAT 564
111 111 111 111 111 111 111 111 111 111
275 spGluLeuLeuGluSerAlaArgSerGlyAsnGluGluLysMetMetAla 291
565 ATCTCTCTGTACTACAAGCATCCACAACATCCAAACACTGAGGGTAA 614
111 111 111 111 111 111 111 111 111 111
292 LeuLeuThrProLeuAsnValAsnCysHisAlaSerAspGlyArgLysSe 308
615 CACTCTCTACACTTAGCTGTGTAGGAGAGAGTGAAGAAGCAAAAC 664
111 111 111 111 111 111 111 111 111 111
308 rThrProLeuHisLeuAlaAlaGlyTyrAsnArgValLysIleValGlnL 325
665 TGCTGGTGTCCCAAGGAGCAAGTATTACATTGAGATTAAGAAGAAAG 714
111 111 111 111 111 111 111 111 111 111
325 euLeuLeuGlnHisGlyAlaAspValHisAlaLysAspLysGlyAspLeu 341
715 ACACCCCTGCAAGTGGCC 732
111 111 111 111 111 111 111 111 111 111
```


XX 29-JUN-1999; 99US-0141582.
XX (ICOS-) ICOS CORP.
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX DR N-PSDB; AAF63953.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX PT inflammatory and autoimmune disorders -
XX PS Claim 3; Page 200-203; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
XX CC human tankyrase2 protein. This is found in two different versions,
XX CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX CC polyADP-ribosylation activity and is involved in the modification of
XX CC TRF1, which is a telomere-specific binding protein. The regulation of
XX CC telomere length, in which TRF1 has a role, is linked to ageing and
XX CC cancer. The sequences are useful in the treatment of cancers and
XX CC inflammatory disorders.
XX SQ Sequence 1166 AA;

alignment_scores:
Quality: 293-50 Length: 253
Ratio: 1-823 Gaps: 4
Percent Similarity: 63.636 Percent Identity: 31.621
alignment_block:
US-09-509-775-1 x AAB66295 ..
Align seg 1/1 to: AAB66295 from: 1 to: 1166
40 AGCCGGCAGCTGAGGGGGTGTCTGCGCGCAGCAAGTAGTTCTGCTGGG 89
||| ||||| ||||| ||| ||| |||||
2 SerGlyArgArgCysAlaGlyGlyAlaAlaCysAlaSerAlaAla 18
90 CAGCGAAATGAGGGGTGTGTCTAACCTAATG...GTCTGCAACCTGG 136
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18 aGluValGluProAlaAlaArgGluLeuPheGluAlaCysArg... 33
137 CCTCAGCGGAAGCTGGAAGAGTGAAGGAGTAGTATTCTGGCGGATAAA 186
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34AsnGlyAspValGluArgValGlyArgLeuValThrProGluLys 48
187 TCCCTGGCTACTAGACTGACCAGCAGCAGCAAGTGCATGCTGCTGGC 236
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 ValAsnSerArgAspThrAlaGlyArgLysSerThrProLeuHisPheAl 65
237 ATGCTCAGCTGGACATACAGAAATGTTGAATTTTCTGCAACTGGAG 286
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 aAlaGlyPheGlyArgLysAspValValGluThrLeuGlnAsnGlyA 82
287 TGCCAGTGAATGAAGAAGCAGTCAGGTGGTCTCTCTTCATATTGCG 336
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 laAsnValGlnAlaArgAspGlyGlyLeuLeuProLeuHisAsnAla 98
337 GCTTCTGCTGGCGGATGAGATTGTAAGCCCTCTGCGGAAGAGTGC 386
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 CysSerPheGlyHisAlaGluValAlaLeuLeuLeuLeuArgHisGlyAl 115
387 TCAAGTGAATGCTGTCTCAATCAAAATGGCTGCTACTCCCTTACATTATGCG 436
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 aAspProAsnAlaArgAspAsnTrpAsnThrProLeuHisGluAla 132
437 CTTCGAAACAGGCATGAGATCGCTGCTCATGTTACTGGAGCGGGCT 486
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 laLeuGlyGlyLysLeuAspValCysLeuValLeuLeuGlnHisGlyAla 148

487 AATCCAGATGCTAAGGAC 504
::: |||
149 GluProThrIleArgAsnThrAspGlyArgThrAlaLeuAspLeuAlaAs 165
505CATTATGAGGCTACAGCAATGC 526
::: |||
165 pProSerAlaLysAlaValLeuThrGlyGluTyrLysLysAspGluLeuL 182
527 ACCGGGAGAGCCAGGGTAAC...TTGAAGATGATTATCATCTCTTCTG 573
::: |||
182 euGluSerAlaArgSerGlyAsnGluGluLysMetMetAlaLeuLeuThr 198
574 TACTACAAAGCATCCCAACATCCCAAGACACTCAGGGTAACACTCTCT 623
::: |||
199 ProLeuAsnValAsnCysHisAlaSerAspGlyArgLysSerThrProLe 215
624 ACATTAGCTGTGATGAGGAGAGAGTGAAGAGCAAAACTGCTGTGT 673
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 uHisLeuAlaAlaGlyTyrAsnArgValLysIleValGlnLeuLeuLeuG 232
674 CCCAAGGAGCAAGTATTACATTGAGATAAAGAAAGAACACACCCCTG 723
::: |||
232 InHisGlyAlaAspValHisAlaLysAspLysGlyAspLeuValProLeu 248
724 CAAGTGGCC 732
::: |||
249 HisAsnAla 251

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT: AAB47023

seq_documentation_block:

ID AAB47023 standard; Protein; 991 AA.

XX AAB47023;

XX DT 29-MAR-2001 (first entry)

XX DE Mouse SPANK.

XX KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.

XX OS Mus musculus.

XX FT Key Location/Qualifiers

XX FT Misc-difference 483 /note= "Encoded by G"

XX FT Misc-difference 614..623 /note= "Encoded by G"

XX FT Misc-difference 776 /note= "Encoded by ATGCC"

XX PN WO200077225-A1.

XX PD 21-DEC-2000.

XX PF 09-JUN-2000; 2000WO-US15926.

XX PR 11-JUN-1999; 99US-0138957.

XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX PA (GEOH) GEN HOSPITAL CORP.

XX PI Chi N, Lodish HF;

XX WP; 2001-091404/10.

XX DR

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 32.29 Seconds
(without alignments)
424.312 Million cell updates/sec

Title: US-09-509-775-2

Perfect score: 1164

Sequence: 1 MEGCVSNLWNCNLAYSCKLE.....TPLQVAKGGLGLIKRWVEG 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1164	100.0	226	AAW15483	Human P28. Homo s
2	1164	100.0	226	AAW02430	Human gankyrin pro
3	1110	95.4	231	AAW02432	Rat gankyrin prote
4	1105	94.9	231	AAW02431	Mouse gankyrin pro
5	288.5	24.8	1166	AAW47022	Human SPANK. Homo
6	287	24.7	352	AAW11616	D. immitis ankyrin
7	287	24.7	1745	AAW70608	Full length ankyri
8	287	24.7	1745	AAW76776	D. immitis ankyrin
9	287	24.7	1745	AAW11589	D. immitis ankyrin
10	283.5	24.4	522	AAW66287	Human tankyrase2 c
11	283.5	24.4	1166	AAW66295	Human tankyrase2 T

12	283.5	24.4	1169	22	AAW66278	Human tankyrase2 r
13	283.5	24.4	1169	22	AAW66288	Human tankyrase2 c
14	283.5	24.4	1262	22	AAW66290	Human tankyrase2 c
15	283.5	24.4	1385	22	AAW66294	Human tankyrase2 T
16	281.5	24.2	673	21	AAW44403	Human truncated ta
17	281.5	24.2	949	21	AAW44404	Human SPANK. Mus
18	281.5	24.2	991	22	AAW47023	Mouse SPANK. Mus
19	281.5	24.2	1327	21	AAW27212	Human tankyrase I
20	281.5	24.2	1327	21	AAW44402	Human tankyrase.
21	281.5	24.2	1327	22	AAW66279	Human tankyrase2 S
22	281	24.1	1181	22	AAW66297	Drosophila tankyrase
23	280.5	24.1	1166	22	AAW72589	Human tankyrase ho
24	277.5	23.8	1166	21	AAW27211	Human tankyrase II
25	271	23.3	302	19	AAW70609	Ankyrin protein PB
26	271	23.3	302	19	AAW76777	B. malayi ankyrin
27	271	23.3	302	21	AAW11590	B. malaya ankyrin
28	270.5	23.2	1030	19	AAW53572	Human myosin light
29	270.5	23.2	1030	19	AAW41378	Human protein p164
30	269.5	23.2	976	19	AAW53571	Rat p138 protein.
31	269.5	23.2	976	19	AAW41377	Rat protein p138.
32	269	23.1	1074	20	AAW05734	Human Grb7 effecto
33	265	22.8	756	22	AAW66286	Human tankyrase2 c
34	265	22.8	784	22	AAW66285	Human tankyrase2 c
35	265	22.8	907	22	AAW48574	Human breast cance
36	262	22.5	303	19	AAW70606	Ankyrin protein fr
37	262	22.5	303	19	AAW76774	D. immitis ankyrin
38	262	22.5	303	21	AAW11587	D. immitis ankyrin
39	260.5	22.4	978	21	AAW42288	Human ORFX ORF2052
40	258.5	22.2	763	21	AAW79154	Mouse protein kina
41	258.5	22.2	786	21	AAW69163	Amino acid sequenc
42	258.5	22.2	787	21	AAW76079	Murine protein kin
43	258.5	22.2	787	22	AAW56018	Skin cell protein,
44	252.5	21.7	456	21	AAW12893	Arabidopsis thalia
45	252.5	21.7	456	21	AAW27402	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW15483
ID AAW15483 standard; Protein; 226 AA.
XX
AC AAW15483;
XX
XX
DT 17-JUN-1997 (first entry)
XX
DE Human P28.
XX
KW Human; proteasome; P28; diagnosis; malignant tumour.
XX
OS Homo sapiens.
XX
PN JP09075085-A.
XX
PD 25-MAR-1997.
XX
PF 13-SEP-1995; 95JP-0235052.
XX
PR 13-SEP-1995; 95JP-0235052.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
XX
DR WPI; 1997-239267/22.
DR N-PSDB; AAT66424-25.
XX
PT Human 26S proteasome constituting component protein - useful in the
PT diagnosis of e.g. malignant tumour
XX
PS Claim 1; Page 6-7; 9pp; Japanese.
XX
CC This sequence represents the human proteasome component protein p28.
CC The protein, p28, is useful for the diagnosis and treatment of

CC various diseases caused by proteasomes such as malignant tumour.

SQ Sequence 226 AA; Query Match 100.0%; Score 1164; DB 18; Length 226;
Best Local Similarity 100.0%; Pred. No. 5e-119;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
|||||
Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
|||||

QY 61 QLGVPVNDKDDAGWSPHLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHE 120
|||||
Db 61 qlgvpvndkddagwsphlhlaasagrdeivkallgkgagvnavngctplhyaasknrhe 120
|||||

QY 121 IAVMLLEGANPDADKDYEAATAMHRAAAGNLKMIHILLYKASTNIQDTGNTPLHLAC 180
|||||
Db 121 iavmlleganpdadkdyeatamhraaagnlkmihillykastniqdtgntplhlac 180
|||||

QY 181 DEERVEEAKLLVSGASIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
|||||
Db 181 deerveeakllvsgasiyienkeektplqvakggglilkrmvveg 226
|||||

RESULT 2

AAAY02430
ID AAY02430 standard; Protein; 226 AA.
XX AC AAY02430;
XX DE Human gankyrin protein.
XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
XX KW hepatocellular carcinoma; oncogenesis mechanism.
XX OS Homo sapiens.
XX PN WO9918201-A1.
XX PD 15-APR-1999.
XX PF 02-OCT-1998; 98WO-JP04467.
XX PR 03-OCT-1997; 97JP-0286214.
XX PA (FUJITA) FUJITA.
XX PI Fujita J;
XX DR WPI; 1999-277266/23.
XX DR N-PSDB; AAX35852.
XX PT Gankyrin polypeptides, useful for treatment and diagnosis of
XX PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
XX PT mechanism
XX PS Claim 1; Page 70-71; 111pp; Japanese.
XX CC The specification describes human, murine and rat gankyrin DNA and
XX CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
XX CC ability and apoptosis induction. The polypeptides and their antibodies
XX CC can be used in the diagnosis and treatment of cancers,
XX CC e.g. hepatocellular carcinoma, and study of oncogenesis
XX CC mechanism.
XX CC The present sequence represents human gankyrin.
XX SQ Sequence 226 AA;

Query Match

100.0%; Score 1164; DB 20; Length 226;

Best Local Similarity 100.0%; Pred. No. 5e-119;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
|||||
Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
|||||

QY 61 QLGVPVNDKDDAGWSPHLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHE 120
|||||
Db 61 qlgvpvndkddagwsphlhlaasagrdeivkallgkgagvnavngctplhyaasknrhe 120
|||||

QY 121 IAVMLLEGANPDADKDYEAATAMHRAAAGNLKMIHILLYKASTNIQDTGNTPLHLAC 180
|||||
Db 121 iavmlleganpdadkdyeatamhraaagnlkmihillykastniqdtgntplhlac 180
|||||

QY 181 DEERVEEAKLLVSGASIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
|||||
Db 181 deerveeakllvsgasiyienkeektplqvakggglilkrmvveg 226
|||||

RESULT 3

AAAY02432
ID AAY02432 standard; Protein; 231 AA.
XX AC AAY02432;
XX DT 14-JUL-1999 (first entry)
XX DE Rat gankyrin protein.
XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
XX KW hepatocellular carcinoma; oncogenesis mechanism.
XX OS Rattus sp.
XX PN WO9918201-A1.
XX PD 15-APR-1999.
XX PF 02-OCT-1998; 98WO-JP04467.
XX PR 03-OCT-1997; 97JP-0286214.
XX PA (FUJITA) FUJITA.
XX PI Fujita J;
XX DR WPI; 1999-277266/23.
XX DR N-PSDB; AAX35854.
XX PT Gankyrin polypeptides, useful for treatment and diagnosis of
XX PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
XX PT mechanism
XX PS Claim 1; Page 76-78; 111pp; Japanese.
XX CC The specification describes human, murine and rat gankyrin DNA and
XX CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
XX CC ability and apoptosis induction. The polypeptides and their antibodies
XX CC can be used in the diagnosis and treatment of cancers,
XX CC e.g. hepatocellular carcinoma, and study of oncogenesis
XX CC mechanism.
XX CC The present sequence represents rat gankyrin.
XX SQ Sequence 231 AA;

Query Match

95.4%; Score 1110; DB 20; Length 231;
Best Local Similarity 95.1%; Pred. No. 4.1e-113;
Matches 214; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
|||||
Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
|||||

	Query Match	94.9%;	Score 1105;	DB 20;	Length 231;
	Best Local Similarity	93.3%;	Pred. No. 1.4e-112;		
	Matches 210;	Conservative 10;	Mismatches 5;	Indels 0;	Gaps
QY	1	MEGCVSNLWCNLAYSGKLEELKESILADKSLSATRTDQDSRTLHWCSAGHTEIVEFILL	60		
		: : : :			
Db	1	megcvsnimicnlaysgkideikeriladkslatrtddgsrtallhwcsaghteivefill	60		
		: : : :			
QY	61	QLGVPNVDKDDAGSPHLHAASAGRDEIVKALLKGGAOVNAVNONGCCTPLHYAASKNRHE	120		
		: : : :			
Db	61	qlgvpnvdkdagwsphlhaasagrdeivkalivkgahvsnvnqngcctplhyaaaknrhe	120		
		: : : :			
QY	121	IAYMVLLEGANPDOKDHYEATAMHRAAAAKGLKMTHLLYYRKASTNIQDTEGNTPHLIAC	180		

AA New insulin signalling protein SPANK, useful for reducing body mass,
 PT glucose intolerance or insulin resistance and for preventing or
 PT treating obesity-related and muscle-related diseases -
 XX
 XX
 PS Claim 3; Fig 3; 65pp; English.
 XX
 CC This sequence represents human SPANK. The SPANK protein comprises
 CC 3 domains:
 CC (a) a SAM (sterile alpha motif) domain;
 CC (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
 CC domain; and
 CC (c) an ANK domain composed of ankyrin repeats.
 CC SPANK is a cytosolic protein which can poly(ADP-ribosyl)ate itself.
 CC SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
 CC translocation of GLUT4 in the perinuclear region of adipocytes. It
 CC is an effector in the insulin signalling pathway in eukaryotic cells
 CC

PA (HESK-) HESKA CORP.

XX Blehm ES, Tang L;

XX WPI: 2000-375493/32.

DR N-PSDB; AAA58193, AAA58194, AAA58195, AAA58196.

XX New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding
PT them, useful for treating and protecting animals from diseases caused
PT by parasitic helminths, e.g. heartworm disease, elephantiasis or
PT hydrocele

XX Example 1; Column 117-128; 120pp; English.

XX The invention relates to ankyrin proteins and nucleic acids from the
CC parasitic helminths Dirofilaria immitis and Brugia malayi. It also
CC relates to antibodies raised against such ankyrin proteins and to
CC compounds that inhibit Dirofilaria or Brugia ankyrin function.
CC Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour
CC L3 cDNA library using PCR primers based on the sequence of the E1
CC ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin
CC UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult
CC female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR
CC primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids
CC represent novel targets for anti-helminthic vaccines and drugs. Ankyrin
CC nucleic acid molecules, proteins, vaccines and compositions are useful
CC for protecting animals, particularly dogs, from diseases caused by
CC parasitic helminths (e.g., heartworm disease, elephantiasis or
CC hydrocele), as well as for treating the infection. The ankyrin nucleic
CC acid molecules, proteins, vaccines and compositions of the invention are
CC especially useful in treating and preventing infections caused by
CC nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria,
CC strongylid, strongyloides, trichostrongyle, or trichurid nematodes and
CC are also useful against cestodes and trematodes. The therapeutic
CC compositions may be administered to mammals, including dogs, cats,
CC humans, ferrets, horses, cattle, sheep, and other pets; economic food
CC animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and
CC compounds may also be used as diagnostic reagents to detect infection by
CC parasitic helminths. Prior art anti-helminthic drugs require repeated
CC administration, which often leads to the development of resistant
CC helminth strains that no longer respond to treatment. Such drugs can
CC also cause harmful side effects in the individual being treated, and a
CC number of these drugs can only treat the symptoms of a parasitic disease,
CC being unable to prevent infection by the parasitic helminth. Elucidation
CC of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates
CC the development of agents which inhibit ankyrin-mediated parasite
CC developmental and migratory pathways. Sequence AAB11589 represents full-
CC length D. immitis ankyrin, and sequences AAB11582-B11588, AAB11591-B11614
CC and AAB11616-B11629 represent D. immitis ankyrin fragments.

XX Sequence 1745 AA;

Query Match 24.7%; Score 287; DB 21; Length 1745;

Best Local Similarity 36.5%; Pred. No. 7.7e-22;

Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 20 BELKESILAE-----DKSLATRTDQDSRTALHWACSGHTEIVEFLLQLGVPVNDKDDA 72

DB 521 kegqeevaalmdhgtktiltk---kgftplhaaaygnlpvaksllertgtpydiegkn 577

QY 73 GWSPLHTAASGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANP 132

DB 578 qvtplhvaahyndnkvalillengasahaakngytpplhaaakngmdastillhykana 637

QY 133 DAKDHYEATAMHRAAKGNLKMIIHLLYYKASTNIQTEGNTPLHLACDERVEEAKLLV 192

DB 638 naeskagftplhlaaqeghremaaillengakvgagangrltpmhlcaqedrsvaeely 697

QY 193 SQGASIVYENKEETPLQVA 212

DB 698 kenaaipktkagtytplhva 717

RESULT 10

AAB66287

ID AAB66287 standard; Protein; 522 AA.

XX AAB66287;

XX 05-APR-2001 (first entry)

DE Human tankyrase2 clone protein sequence SEQ ID NO: 99.

KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI: 2001-102896/11.

XX N-PSDB; AAF63925.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX Example 1; Page 156-157; 242pp; English.

XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.

XX Sequence 522 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 522;

Best Local Similarity 32.1%; Pred. No. 3e-22;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 16 SKLEELKESILADKSLATRTDQDSRTALHWACSGHTEIVEFLLQLGVPVNDKDDAGWS 75

DB 37 ngdvervkrilvtpekvnsrdtagrkstplhfaagfgrkdvveyllnganvqarddglli 96

QY 76 PLHTAASAGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDAK 135

DB 97 plhnacsfghaeavnlllrhgadpnrdrwnytpplheaaikgkidvcivilqhgaeptr 156

QY 136 D-----HYTEATAMHRAAKGN-LKMIHLLYYKASTNIQDTGENT 174

DB 157 ntgdrtaldladpsakavltgkdkdelllesarsgnekmmalltpinvnchadgrkst 216

QY 175 PLHLACDEEVEEAKLLVSOGASIIYENKEETPLQVA 212

DB 217 plhlaagynrvkvqlllqhgdvhakdkgdlvplhna 254

RESULT 11

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

XX AAB66295;
AC
XX
DT
XX
XX
DE
XX
KW
KW
XX
XX
OS
XX
PN
XX
PD
XX
XX
PF
XX
PR
XX
PA
XX
XX
PI
XX
DR
DR
DR
XX
XX
PT
PT
PT
XX
PS
XX
XX
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

05-APR-2001 (first entry)
Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.
Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
inflammatory disorder.
Homo sapiens.
WO200100849-A1.
04-JAN-2001.
28-JUN-2000; 2000WO-US17827.
29-JUN-1999; 99US-0141582.
(ICOS-) ICOS CORP.
Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
WPI; 2001-102896/11.
N-PSDB; AAF63953.
New tankyrase2 polypeptides, useful for treating conditions mediated by
poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
inflammatory and autoimmune disorders -
Claim 3; Page 200-203; 242pp; English.
The present invention provides the protein and coding sequence for the
human tankyrase2 protein. This is found in two different versions,
designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
polyADP-ribosylation activity and is involved in the modification of
TRF1, which is a telomere-specific binding protein. The regulation of
telomere length, in which TRF1 has a role, is linked to ageing and
cancer. The sequences are useful in the treatment of cancers and
inflammatory disorders.
Sequence 1166 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1166;
Best Local Similarity 32.1%; Pred. No. 1e-21;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

Qy 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLGVVNDKDDAGWS 75
Db 34 ngdvervrlvtpekvnsrdtagrktstplhfaagfgrkdvveyllqnganvqarddggli 93
Qy 76 PLHTAASAGRDEIVKALGKGAQVNAVNONGCTPLHYAASKNRHIEIAYMLLEGGANPDAK 135
Db 94 plhnacsfghaevnllrhgdpnrdnwnytpiheaaikgkdvciqlhgaeptr 153
Qy 136 D-----HYEATAMHRAAAGN-LKMIHILLYKASTNIQDTEGT 174
Db 154 ntdgrtalldapsakavltgeykkdelllesarsgneekmmalltplnvchnasgrkst 213
Qy 175 PLHACDEERVEEAKLLVSGASIVYENKEETPLQVA 212
Db 214 plhaagynrvkivqlhghadghakdgdlvplhna 251

RESULT 12
AAB66278
ID AAB66278 standard; Protein; 1169 AA.
XX
AC AAB66278;
XX
DT 05-APR-2001 (first entry)

XX Human tankyrase2 related protein sequence SEQ ID NO: 2.
XX
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US17827.
XX
XX 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
XX WPI; 2001-102896/11.
XX
XX N-PSDB; AAF63837.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX
XX Disclosure; Page 109-113; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.
XX
XX Sequence 1169 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1169;
Best Local Similarity 32.1%; Pred. No. 1e-21;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

Qy 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLQLGVVNDKDDAGWS 75
Db 37 ngdvervrlvtpekvnsrdtagrktstplhfaagfgrkdvveyllqnganvqarddggli 96
Qy 76 PLHTAASAGRDEIVKALGKGAQVNAVNONGCTPLHYAASKNRHIEIAYMLLEGGANPDAK 135
Db 97 plhnacsfghaevnllrhgdpnrdnwnytpiheaaikgkdvciqlhgaeptr 156
Qy 136 D-----HYEATAMHRAAAGN-LKMIHILLYKASTNIQDTEGT 174
Db 157 ntdgrtalldapsakavltgeykkdelllesarsgneekmmalltplnvchnasgrkst 216
Qy 175 PLHACDEERVEEAKLLVSGASIVYENKEETPLQVA 212
Db 217 plhaagynrvkivqlhghadghakdgdlvplhna 254

RESULT 13
AAB66288
ID AAB66288 standard; Protein; 1169 AA.
XX
XX AAB66288;
XX
XX
DT 05-APR-2001 (first entry)
XX
XX Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder.

OS Homo sapiens.

PN WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

XX N-PSDB; AAF63926.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

XX Example 1; Page 162-1665; 242pp; English.

XX The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

XX Sequence 1169 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1169;

Best Local Similarity 32.1%; Pred. No. 1e-21;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWS 75

Db 37 ngdverkrivtpkvnrdtagrkstplhfaagfgrkdvveyllnganvqarddggli 96

QY 76 PLHTASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAK 135

Db 97 plhnacsfghaevnnllrhgdpnrdwnnytpplheaalkgkdvivllghgaeptr 156

QY 136 D-----HVEATAMHRAAKGN-LKMIHILLYYKASTNIQDTGNT 174

Db 157 ntgdrtaldladsakavltgkykdellesarsgnekmmalltplnvchnasdrkst 216

QY 175 PLHLACDEERVEAKLLVSOGASIYIENKEETPLQVA 212

Db 217 plhlaagynrvkivqlllghgadvhakdkgdlvplhna 254

RESULT 14

AAB66290

ID AAB66290 standard; Protein; 1262 AA.

XX AAB66290;

XX 05-APR-2001 (first entry)

XX Human tankyrase2 clone consensus protein SEQ ID NO: 107.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX inflammatory disorder.

XX Homo sapiens.

XX

PN WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

XX N-PSDB; AAF63930.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

XX Example 2; Page 173-176; 242pp; English.

XX The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

XX Sequence 1262 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1262;

Best Local Similarity 32.1%; Pred. No. 1.1e-21;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWS 75

Db 130 ngdverkrivtpkvnrdtagrkstplhfaagfgrkdvveyllnganvqarddggli 189

QY 76 PLHTASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAK 135

Db 190 plhnacsfghaevnnllrhgdpnrdwnnytpplheaalkgkdvivllghgaeptr 249

QY 136 D-----HVEATAMHRAAKGN-LKMIHILLYYKASTNIQDTGNT 174

Db 250 ntgdrtaldladsakavltgkykdellesarsgnekmmalltplnvchnasdrkst 309

QY 175 PLHLACDEERVEAKLLVSOGASIYIENKEETPLQVA 212

Db 310 plhlaagynrvkivqlllghgadvhakdkgdlvplhna 347

RESULT 15

AAB66294

ID AAB66294 standard; Protein; 1385 AA.

XX AAB66294;

XX 05-APR-2001 (first entry)

XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX inflammatory disorder.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX

THIS PAGE BLANK (USPTO)

CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z26634; CAB42644.1; -;
DR EMBL; X56957; CAA40278.1; -;
DR EMBL; X56958; CAA40279.2; -;
DR EMBL; M37123; AAA62828.1; -;
DR PIR; S14533; S14533.
DR PIR; A39643; A39643.
DR PIR; B39643; B39643.
DR PIR; S14569; S14569.
DR HSP; Q00420; IAWC.
DR MIM; 106410; -;
DR InterPro; IPR000488; -;
DR InterPro; IPR000906; -;
DR InterPro; IPR002110; -;
DR Pfam; PF00791; ZUS; 1.
DR Pfam; PF00531; ank; 22.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation; Multigene family.
FT REPEAT 63 92
FT REPEAT 96 125
FT REPEAT 129 158
FT REPEAT 162 191
FT REPEAT 193 220
FT REPEAT 232 261
FT REPEAT 265 294
FT REPEAT 298 327
FT REPEAT 331 360
FT REPEAT 364 393
FT REPEAT 397 426
FT REPEAT 430 459
FT REPEAT 463 492
FT REPEAT 496 525
FT REPEAT 529 558
FT REPEAT 562 591
FT REPEAT 595 624
FT REPEAT 628 657
FT REPEAT 661 690
FT REPEAT 694 723
FT REPEAT 727 756
FT REPEAT 760 789
FT REPEAT 793 822
FT REPEAT 825 854
FT REPEAT 857 886
FT REPEAT 889 918
FT REPEAT 921 950
FT REPEAT 953 982
FT REPEAT 985 1014
FT REPEAT 1017 1046
FT REPEAT 1049 1078
FT REPEAT 1081 1110
FT REPEAT 1113 1142
FT REPEAT 1145 1174
FT REPEAT 1177 1206
FT REPEAT 1209 1238
FT REPEAT 1241 1270
FT REPEAT 1273 1302
FT REPEAT 1305 1334
FT REPEAT 1337 1366
FT REPEAT 1369 1398
FT REPEAT 1401 1430
FT REPEAT 1433 1462
FT REPEAT 1465 1494
FT REPEAT 1497 1526
FT REPEAT 1529 1558
FT REPEAT 1561 1590
FT REPEAT 1593 1622
FT REPEAT 1625 1654
FT REPEAT 1657 1686
FT REPEAT 1689 1718
FT REPEAT 1721 1750
FT REPEAT 1753 1782
FT REPEAT 1785 1814
FT REPEAT 1817 1846
FT REPEAT 1849 1878
FT REPEAT 1881 1910
FT REPEAT 1913 1942
FT REPEAT 1945 1974
FT REPEAT 1977 2006
FT REPEAT 2009 2038
FT REPEAT 2041 2070
FT REPEAT 2073 2102
FT REPEAT 2105 2134
FT REPEAT 2137 2166
FT REPEAT 2169 2198
FT REPEAT 2201 2230
FT REPEAT 2233 2262
FT REPEAT 2265 2294
FT REPEAT 2297 2326
FT REPEAT 2329 2358
FT REPEAT 2361 2390
FT REPEAT 2393 2422
FT REPEAT 2425 2454
FT REPEAT 2457 2486
FT REPEAT 2489 2518
FT REPEAT 2521 2550
FT REPEAT 2553 2582
FT REPEAT 2585 2614
FT REPEAT 2617 2646
FT REPEAT 2649 2678
FT REPEAT 2681 2710
FT REPEAT 2713 2742
FT REPEAT 2745 2774
FT REPEAT 2777 2806
FT REPEAT 2809 2838
FT REPEAT 2841 2870
FT REPEAT 2873 2902
FT REPEAT 2905 2934
FT REPEAT 2937 2966
FT REPEAT 2969 2998
FT REPEAT 3001 3030
FT REPEAT 3033 3062
FT REPEAT 3065 3094
FT REPEAT 3097 3126
FT REPEAT 3129 3158
FT REPEAT 3161 3190
FT REPEAT 3193 3222
FT REPEAT 3225 3254
FT REPEAT 3257 3286
FT REPEAT 3289 3318
FT REPEAT 3321 3350
FT REPEAT 3353 3382
FT REPEAT 3385 3414
FT REPEAT 3417 3446
FT REPEAT 3449 3478
FT REPEAT 3481 3510
FT REPEAT 3513 3542
FT REPEAT 3545 3574
FT REPEAT 3577 3606
FT REPEAT 3609 3638
FT REPEAT 3641 3670
FT REPEAT 3673 3702
FT REPEAT 3705 3734
FT REPEAT 3737 3766
FT REPEAT 3769 3798
FT REPEAT 3801 3830
FT REPEAT 3833 3862
FT REPEAT 3865 3894
FT REPEAT 3897 3926
FT REPEAT 3929 3958
FT REPEAT 3961 3990
FT REPEAT 3993 4022
FT REPEAT 4025 4054
FT REPEAT 4057 4086
FT REPEAT 4089 4118
FT REPEAT 4121 4150
FT REPEAT 4153 4182
FT REPEAT 4185 4214
FT REPEAT 4217 4246
FT REPEAT 4249 4278
FT REPEAT 4281 4310
FT REPEAT 4313 4342
FT REPEAT 4345 4374
FT REPEAT 4377 4406
FT REPEAT 4409 4438
FT REPEAT 4441 4470
FT REPEAT 4473 4502
FT REPEAT 4505 4534
FT REPEAT 4537 4566
FT REPEAT 4569 4598
FT REPEAT 4601 4630
FT REPEAT 4633 4662
FT REPEAT 4665 4694
FT REPEAT 4697 4726
FT REPEAT 4729 4758
FT REPEAT 4761 4790
FT REPEAT 4793 4822
FT REPEAT 4825 4854
FT REPEAT 4857 4886
FT REPEAT 4889 4918
FT REPEAT 4921 4950
FT REPEAT 4953 4982
FT REPEAT 4985 5014
FT REPEAT 5017 5046
FT REPEAT 5049 5078
FT REPEAT 5081 5110
FT REPEAT 5113 5142
FT REPEAT 5145 5174
FT REPEAT 5177 5206
FT REPEAT 5209 5238
FT REPEAT 5241 5270
FT REPEAT 5273 5302
FT REPEAT 5305 5334
FT REPEAT 5337 5366
FT REPEAT 5369 5398
FT REPEAT 5401 5430
FT REPEAT 5433 5462
FT REPEAT 5465 5494
FT REPEAT 5497 5526
FT REPEAT 5529 5558
FT REPEAT 5561 5590
FT REPEAT 5593 5622
FT REPEAT 5625 5654
FT REPEAT 5657 5686
FT REPEAT 5689 5718
FT REPEAT 5721 5750
FT REPEAT 5753 5782
FT REPEAT 5785 5814
FT REPEAT 5817 5846
FT REPEAT 5849 5878
FT REPEAT 5881 5910
FT REPEAT 5913 5942
FT REPEAT 5945 5974
FT REPEAT 5977 6006
FT REPEAT 6009 6038
FT REPEAT 6041 6070
FT REPEAT 6073 6102
FT REPEAT 6105 6134
FT REPEAT 6137 6166
FT REPEAT 6169 6198
FT REPEAT 6201 6230
FT REPEAT 6233 6262
FT REPEAT 6265 6294
FT REPEAT 6297 6326
FT REPEAT 6329 6358
FT REPEAT 6361 6390
FT REPEAT 6393 6422
FT REPEAT 6425 6454
FT REPEAT 6457 6486
FT REPEAT 6489 6518
FT REPEAT 6521 6550
FT REPEAT 6553 6582
FT REPEAT 6585 6614
FT REPEAT 6617 6646
FT REPEAT 6649 6678
FT REPEAT 6681 6710
FT REPEAT 6713 6742
FT REPEAT 6745 6774
FT REPEAT 6777 6806
FT REPEAT 6809 6838
FT REPEAT 6841 6870
FT REPEAT 6873 6902
FT REPEAT 6905 6934
FT REPEAT 6937 6966
FT REPEAT 6969 6998
FT REPEAT 7001 7030
FT REPEAT 7033 7062
FT REPEAT 7065 7094
FT REPEAT 7097 7126
FT REPEAT 7129 7158
FT REPEAT 7161 7190
FT REPEAT 7193 7222
FT REPEAT 7225 7254
FT REPEAT 7257 7286
FT REPEAT 7289 7318
FT REPEAT 7321 7350
FT REPEAT 7353 7382
FT REPEAT 7385 7414
FT REPEAT 7417 7446
FT REPEAT 7449 7478
FT REPEAT 7481 7510
FT REPEAT 7513 7542
FT REPEAT 7545 7574
FT REPEAT 7577 7606
FT REPEAT 7609 7638
FT REPEAT 7641 7670
FT REPEAT 7673 7702
FT REPEAT 7705 7734
FT REPEAT 7737 7766
FT REPEAT 7769 7798
FT REPEAT 7801 7830
FT REPEAT 7833 7862
FT REPEAT 7865 7894
FT REPEAT 7897 7926
FT REPEAT 7929 7958
FT REPEAT 7961 7990
FT REPEAT 7993 8022
FT REPEAT 8025 8054
FT REPEAT 8057 8086
FT REPEAT 8089 8118
FT REPEAT 8121 8150
FT REPEAT 8153 8182
FT REPEAT 8185 8214
FT REPEAT 8217 8246
FT REPEAT 8249 8278
FT REPEAT 8281 8310
FT REPEAT 8313 8342
FT REPEAT 8345 8374
FT REPEAT 8377 8406
FT REPEAT 8409 8438
FT REPEAT 8441 8470
FT REPEAT 8473 8502
FT REPEAT 8505 8534
FT REPEAT 8537 8566
FT REPEAT 8569 8598
FT REPEAT 8601 8630
FT REPEAT 8633 8662
FT REPEAT 8665 8694
FT REPEAT 8697 8726
FT REPEAT 8729 8758
FT REPEAT 8761 8790
FT REPEAT 8793 8822
FT REPEAT 8825 8854
FT REPEAT 8857 8886
FT REPEAT 8889 8918
FT REPEAT 8921 8950
FT REPEAT 8953 8982
FT REPEAT 8985 9014
FT REPEAT 9017 9046
FT REPEAT 9049 9078
FT REPEAT 9081 9110
FT REPEAT 9113 9142
FT REPEAT 9145 9174
FT REPEAT 9177 9206
FT REPEAT 9209 9238
FT REPEAT 9241 9270
FT REPEAT 9273 9302
FT REPEAT 9305 9334
FT REPEAT 9337 9366
FT REPEAT 9369 9398
FT REPEAT 9401 9430
FT REPEAT 9433 9462
FT REPEAT 9465 9494
FT REPEAT 9497 9526
FT REPEAT 9529 9558
FT REPEAT 9561 9590
FT REPEAT 9593 9622
FT REPEAT 9625 9654
FT REPEAT 9657 9686
FT REPEAT 9689 9718
FT REPEAT 9721 9750
FT REPEAT 9753 9782
FT REPEAT 9785 9814
FT REPEAT 9817 9846
FT REPEAT 9849 9878
FT REPEAT 9881 9910
FT REPEAT 9913 9942
FT REPEAT 9945 9974
FT REPEAT 9977 1006
FT REPEAT 1009 1039
FT REPEAT 1041 1070
FT REPEAT 1073 1102
FT REPEAT 1105 1134
FT REPEAT 1137 1166
FT REPEAT 1169 1198
FT REPEAT 1201 1230
FT REPEAT 1233 1262
FT REPEAT 1265 1294
FT REPEAT 1297 1326
FT REPEAT 1329 1358
FT REPEAT 1361 1390
FT REPEAT 1393 1422
FT REPEAT 1425 1454
FT REPEAT 1457 1486
FT REPEAT 1489 1518
FT REPEAT 1521 1550
FT REPEAT 1553 1582
FT REPEAT 1585 1614
FT REPEAT 1617 1646
FT REPEAT 1649 1678
FT REPEAT 1681 1710
FT REPEAT 1713 1742
FT REPEAT 1745 1774
FT REPEAT 1777 1806
FT REPEAT 1809 1838
FT REPEAT 1841 1870
FT REPEAT 1873 1902
FT REPEAT 1905 1934
FT REPEAT 1937 1966
FT REPEAT 1969 1998
FT REPEAT 2001 2030
FT REPEAT 2033 2062
FT REPEAT 2065 2094
FT REPEAT 2097 2126
FT REPEAT 2129 2158
FT REPEAT 2161 2190
FT REPEAT 2193 2222
FT REPEAT 2225 2254
FT REPEAT 2257 2286
FT REPEAT 2289 2318
FT REPEAT 2321 2350
FT REPEAT 2353 2382
FT REPEAT 2385 2414
FT REPEAT 2417 2446
FT REPEAT 2449 2478
FT REPEAT 2481 2510
FT REPEAT 2513 2542
FT REPEAT 2545 2574
FT REPEAT 2577 2606
FT REPEAT 2609 2638
FT REPEAT 2641 2670
FT REPEAT 2673 2702
FT REPEAT 2705 2734
FT REPEAT 2737 2766
FT REPEAT 2769 2798
FT REPEAT 2801 2830
FT REPEAT 2833 2862
FT REPEAT 2865 2894
FT REPEAT 2897 2926
FT REPEAT 2929 2958
FT REPEAT 2961 2990
FT REPEAT 2993 3022
FT REPEAT 3025 3054
FT REPEAT 3057 3086
FT REPEAT 3089 3118
FT REPEAT 3121 3150
FT REPEAT 3153 3182
FT REPEAT 3185 3214
FT REPEAT 3217 3246
FT REPEAT 3249 3278
FT REPEAT 3281 3310
FT REPEAT 3313 3342
FT REPEAT 3345 3374
FT REPEAT 3377 3406
FT REPEAT 3409 3438
FT REPEAT 3441 3470
FT REPEAT 3473 3502
FT REPEAT 3505 3534
FT REPEAT 3537 3566
FT REPEAT 3569 3598
FT REPEAT 3601 3630
FT REPEAT 3633 3662
FT REPEAT 3665 3694
FT REPEAT 3697 3726
FT REPEAT 3729 3758
FT REPEAT 3761 3790
FT REPEAT 3793 3822
FT REPEAT 3825 3854
FT REPEAT 3857 3886
FT REPEAT 3889 3918
FT REPEAT 3921 3950
FT REPEAT 3953 3982
FT REPEAT 3985 4014
FT REPEAT 4017 4046
FT REPEAT 4049 4078
FT REPEAT 4081 4110
FT REPEAT 4113 4142
FT REPEAT 4145 4174
FT REPEAT 4177 4206
FT REPEAT 4209 4238
FT REPEAT 4241 4270
FT REPEAT 4273 4302
FT REPEAT 4305 4334
FT REPEAT 4337 4366
FT REPEAT 4369 4398
FT REPEAT 4401 4430
FT REPEAT 4433 4462
FT REPEAT 4465 4494
FT REPEAT 4497 4526
FT REPEAT 4529 4558
FT REPEAT 4561 4590
FT REPEAT 4593 4622
FT REPEAT 4625 4654
FT REPEAT 4657 4686
FT REPEAT 4689 4718
FT REPEAT 4721 4750
FT REPEAT 4753 4782
FT REPEAT 4785 4814
FT REPEAT 4817 4846
FT REPEAT 4849 4878
FT REPEAT 4881 4910
FT REPEAT 4913 4942
FT REPEAT 4945 4974
FT REPEAT 4977 5006
FT REPEAT 5009 5038
FT REPEAT 5041 5070
FT REPEAT 5073 5102
FT REPEAT 5105 5134
FT REPEAT 5137 5166
FT REPEAT 5169 5198
FT REPEAT 5201 5230
FT REPEAT 5233 5262
FT REPEAT 5265 5294
FT REPEAT 5297 5326
FT REPEAT 5329 5358
FT REPEAT 5361 5390
FT REPEAT 5393 5422
FT REPEAT 5425 5454
FT REPEAT 5457 5486
FT REPEAT 5489 5518
FT REPEAT 5521 5550
FT REPEAT 5553 5582
FT REPEAT 5585 5614
FT REPEAT 5617 5646
FT REPEAT 5649 5678
FT REPEAT 5681 5710
FT REPEAT 5713 5742
FT REPEAT 5745 5774
FT REPEAT 5777 5806
FT REPEAT 5809 5838
FT REPEAT 5841 5870
FT REPEAT 5873 5902
FT REPEAT 5905 5934
FT REPEAT 5937 5966
FT REPEAT 5969 5998
FT REPEAT 6001 6030
FT REPEAT 6033 6062
FT REPEAT 6065 6094
FT REPEAT 6097 6126
FT REPEAT 6129 6158
FT REPEAT 6161 6190
FT REPEAT 6193 6222
FT REPEAT 6225 6254
FT REPEAT 6257 6286
FT REPEAT 6289 6318
FT REPEAT 6321 6350
FT REPEAT 6353 6382
FT REPEAT 6385 6414
FT REPEAT 6417 6446
FT REPEAT 6449 6478
FT REPEAT 6481 6510
FT REPEAT 6513 6542
FT REPEAT 6545 6574
FT REPEAT 6577 6606
FT REPEAT 6609 6638
FT REPEAT 6641 6670
FT REPEAT 6673 6702
FT REPEAT 6705 6734
FT REPEAT 6737 6766
FT REPEAT 6769 6798
FT REPEAT 6801 6830
FT REPEAT 6833 6862
FT REPEAT 6865 6894
FT REPEAT 6897 6926
FT REPEAT 6929 6958
FT REPEAT 6961 6990
FT REPEAT 6993 7022
FT REPEAT 7025 7054
FT REPEAT 7057 7086
FT REPEAT 7089 7118
FT REPEAT 7121 7150
FT REPEAT 7153 7182
FT REPEAT 7185 7214
FT REPEAT 7217 7246
FT REPEAT 7249 7278
FT REPEAT 7281 7310
FT REPEAT 7313 7342
FT REPEAT 7345 7374
FT REPEAT 7377 7406
FT REPEAT 7409 7438
FT REPEAT 7441 7470
FT REPEAT 7473 7502
FT REPEAT 7505 7534
FT REPEAT 7537 7566
FT REPEAT 7569 7598
FT REPEAT 7601 7630
FT REPEAT 7633 7662
FT REPEAT 7665 7694
FT REPEAT 7697 7726
FT REPEAT 7729 7758
FT REPEAT 7761 7790
FT REPEAT 7793 7822
FT REPEAT 7825 7854
FT REPEAT 7857 7886
FT REPEAT 7889 7918
FT REPEAT 7921 7950
FT REPEAT 7953 7982
FT REPEAT 7985 8014
FT REPEAT 8017 8046
FT REPEAT 8049 8078
FT REPEAT 8081 8110
FT REPEAT 8113 8142
FT REPEAT 8145 8174
FT REPEAT 8177 8206
FT REPEAT 8209 8238
FT REPEAT 8241 8270
FT REPEAT 8273 8302
FT REPEAT 8305 8334
FT REPEAT 8337 8366
FT REPEAT 8369 8398
FT REPEAT 8401 8430
FT REPEAT 8433 8462
FT REPEAT 8465 8494
FT REPEAT 8497 8526
FT REPEAT 8529 8558
FT REPEAT 8561 8590
FT REPEAT 8593 8622
FT REPEAT 8625 8654
FT REPEAT 8657 8686
FT REPEAT 8689 8718
FT REPEAT 8721 8750
FT REPEAT 8753 8782
FT REPEAT 8785 8814
FT REPEAT 8817 8846
FT REPEAT 8849 8878
FT REPEAT 8881 8910
FT REPEAT 8913 8942
FT REPEAT 8945 8974
FT REPEAT 8977 9006
FT REPEAT 9009 9038
FT REPEAT 9041 9070
FT REPEAT 9073 9102
FT REPEAT 9105 9134
FT REPEAT 9137 9166
FT REPEAT 9169 9198
FT REPEAT 9201 9230
FT REPEAT 9233 9262
FT REPEAT 9265 9294
FT REPEAT 9297 9326
FT REPEAT 9329 9358
FT REPEAT 9361 9390
FT REPEAT 9393 9422
FT REPEAT 9425 9454
FT REPEAT 9457 9486
FT REPEAT 9489 9518
FT REPEAT 9521 9550
FT REPEAT 9553 9582
FT REPEAT 9585 9614
FT REPEAT 9617 9646
FT REPEAT 9649 9678
FT REPEAT 9681 9710
FT REPEAT 9713 9742
FT REPEAT 9745 9774
FT REPEAT 9777 9806
FT REPEAT 9809 9838
FT REPEAT 9841 9870
FT REPEAT 9873 9902
FT REPEAT 9905 9934
FT REPEAT 9937 9966
FT REPEAT 9969 9998
FT REPEAT 1001 1030
FT REPEAT 1033 1062
FT REPEAT 1065 1094
FT REPEAT 1097 1126
FT REPEAT 1129 1158
FT REPEAT 1161 1190
FT REPEAT 1193 1222
FT REPEAT 1225 1254
FT REPEAT 1257 1286
FT REPEAT 1289 1318
FT REPEAT 1321 1350
FT REPEAT 1353 1382
FT REPEAT 1385 1414
FT REPEAT 1417 1446
FT REPEAT 1449 1478
FT REPEAT 1481 1510
FT REPEAT 1513 1542
FT REPEAT 1545 1574
FT REPEAT 1577 1606
FT REPEAT 1609 1638
FT REPEAT 1641 1670
FT REPEAT 1673 1702
FT REPEAT 1705 1734
FT REPEAT 1737 1766
FT REPEAT 1769 1798
FT REPEAT 1801 1830
FT REPEAT 1833 1862
FT REPEAT 1865 1894
FT REPEAT 1897 1926
FT REPEAT 1929 1958
FT REPEAT 1961 1990
FT REPEAT 1993 2022
FT REPEAT 2025 2054
FT REPEAT 2057 2086
FT REPEAT 2089 2118
FT REPEAT 2121 2150
FT REPEAT 2153 2182
FT REPEAT 2185 2214
FT REPEAT 2217 2246
FT REPEAT 2249 2278
FT REPEAT 2281 2310
FT REPEAT 2313 2342
FT REPEAT 2345 2374
FT REPEAT 2377 2406
FT REPEAT 2409 2438
FT REPEAT 2441 2470
FT REPEAT 2473 2502
FT REPEAT 2505 2534
FT REPEAT 2537 2566
FT REPEAT 2569 2598
FT REPEAT 2601 2630
FT REPEAT 2633 2662
FT REPEAT 2665 2694
FT REPEAT 2697 2726
FT REPEAT 2729 2758
FT REPEAT 2761 2790
FT REPEAT 2793 2822
FT REPEAT 2825 2854
FT REPEAT 2857 2886
FT REPEAT 2889 2918
FT REPEAT 2921 2950
FT REPEAT 2953 2982
FT REPEAT 2985 3014
FT REPEAT 3017 3046
FT REPEAT 3049 3078
FT REPEAT 3081 3110
FT REPEAT 3113 3142
FT REPEAT 3145 3174
FT REPEAT 3177 3206
FT REPEAT 3209 3238
FT REPEAT 3241 3270
FT REPEAT 3273 3302
FT REPEAT 3305 3334
FT REPEAT 3337 3366
FT REPEAT 3369 3398
FT REPEAT 3401 3430
FT REPEAT 3433 3462
FT REPEAT 3465 3494
FT REPEAT 3497 3526
FT REPEAT 3529 3558
FT REPEAT 3561 3590
FT REPEAT 3593 3622
FT REPEAT 3625 3654
FT REPEAT 3657 3686
FT REPEAT 3689 3718
FT REPEAT 3721 3750
FT REPEAT 3753 3782
FT REPEAT 3785 3814
FT REPEAT 3817 3846
FT REPEAT 3849 3878
FT REPEAT 3881 3910
FT REPEAT 3913 3942
FT REPEAT 3945 3974
FT REPEAT 3977 4006
FT REPEAT 4009 4038
FT REPEAT 4041 4070
FT REPEAT 4073 4102
FT REPEAT 4105 4134
FT REPEAT 4137 4166
FT REPEAT 4169 4198
FT REPEAT 4201 4230
FT REPEAT 4233 4262
FT REPEAT 4265 4294
FT REPEAT 4297 4326
FT REPEAT 4329 4358
FT REPEAT 4361 4390
FT REPEAT 4393 4422
FT REPEAT 4425 4454
FT REPEAT 4457 4486
FT REPEAT 4489 4518
FT REPEAT 4521 4550
FT REPEAT 4553 4582
FT REPEAT 4585 4614
FT REPEAT 4617 4646
FT REPEAT 4649 4678
FT REPEAT 4681 4710
FT REPEAT 4713 4742
FT REPEAT 4745 4774
FT REPEAT 4777 4806
FT REPEAT 4809 4838
FT REPEAT 4841 4870
FT REPEAT 4873 4902
FT REPEAT 4905 4934
FT REPEAT 4937 4966
FT REPEAT 4969 4998
FT REPEAT 5001 5030
FT REPEAT 5033 5062
FT REPEAT 5065 5094
FT REPEAT 5097 5126
FT REPEAT 5129 5158
FT REPEAT 5161 5190
FT REPEAT 5193 5222
FT REPEAT 5225 5254
FT REPEAT 5257 5286
FT REPEAT 5289 5318
FT REPEAT 5321 5350
FT REPEAT 5353 5382
FT REPEAT 5385 5414
FT REPEAT 5417 5446
FT REPEAT 5449 5478
FT REPEAT 5481 5510
FT REPEAT 5513 5542
FT REPEAT 5545 5574
FT REPEAT 5577 5606
FT REPEAT 5609 5638
FT REPEAT 5641 5670
FT REPEAT 5673 5702
FT REPEAT 5705 5734
FT REPEAT 5737 5766
FT REPEAT 5769 5798
FT REPEAT 5801 5830
FT REPEAT 5833 5862
FT REPEAT 5865 5894
FT REPEAT 5897 5926
FT REPEAT 5929 5958
FT REPEAT 5961 5990
FT REPEAT 5993 6022
FT REPEAT 6025 6054
FT REPEAT 6057 6086
FT REPEAT 6089 6118
FT REPEAT 6121 6150
FT REPEAT 6153 6182
FT REPEAT 6185 6214
FT REPEAT 6217 6246
FT REPEAT 6249 6278
FT REPEAT 6281 6310
FT REPEAT 6313 6342
FT REPEAT 6345 6374
FT REPEAT 6377 6406
FT REPEAT 6409 6438
FT REPEAT 6441 6470
FT REPEAT 6473 6502
FT REPEAT 6505 6534
FT REPEAT 6537 6566
FT REPEAT 6569 6598
FT REPEAT 6601 6630
FT REPEAT 6633 6662
FT REPEAT 6665 6694
FT REPEAT 6697 6726
FT REPEAT 6729 6758
FT REPEAT 6761 6790
FT REPEAT 6793 6822
FT REPEAT 6825 6854
FT REPEAT 6857 6886
FT REPEAT 6889 6918
FT REPEAT 6921 6950
FT REPEAT 6953 6982
FT REPEAT 6985 7014
FT REPEAT 7017 704


```
FT VARIANT 462 462 /FTid-VAR_000595.
FT V -> I (IN HS).
FT /FTid-VAR_000596.
FT R -> H (IN BRUEGEN).
FT VARIANT 618 618
FT /FTid-VAR_000597.
FT V -> A.
FT VARIANT 749 749
FT /FTid-VAR_000598.
FT D -> E.
FT VARIANT 844 844
FT /FTid-VAR_000599.
FT S -> T.
FT VARIANT 1391 1391
FT /FTid-VAR_000600.
FT E -> D.
FT VARIANT 1285 1285
FT /FTid-VAR_000601.
FT D -> N (IN DUESSELDORF).
FT VARIANT 1591 1591
FT /FTid-VAR_000602.
FT R -> D.
FT VARIANT 1698 1698
FT /FTid-VAR_000603.
FT A -> S (IN REF. 2).
FT CONFLICT 229 229
FT CONFLICT 1545 1545 V -> I (IN REF. 2).
FT SEQUENCE 1880 AA; 206145 MW; 1C5F5E7EFDLCD428 CRC64;

Query Match 24.9%; Score 290; DB 1; Length 1880;
Best Local Similarity 34.3%; Pred. No. 1e-17;
Matches 79; Conservative 38; Mismatches 105; Indels 8; Gaps 3;

QY 5 VSNLMV---CNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGTHIVEFLQ 61
Db 431 VSNVKEVETPLHMAARGHTEVAKYLLONKAKVNAKAKDDQTPHCARIGHTNWKVLLLE 490
QY 62 LGVPVNDKDGWSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEI 121
Db 491 NNANPNLATTAGTHPLHIAAREGHVETVLALEKEASQACMTKGGFTPLHVAAYKGVV 550
QY 122 AVMLLEGANPDADKHVEATAMHRAAKGNLKMHIILLYKASTNTQDTSGNTPHLACD 181
Db 551 AELLERDAHPNAGKNGPLPLHVAHHNLDIVKLLPRGSPHSPAWNGYTPPLHIAK 610
QY 182 EERVEEAKLLVSOGASTYIENKKEKTPLOVA---KGG---LGLILKRMVEG 226
Db 611 ONQVEVARSLLQYGGSANAESVQGVTPPLHIAAQEGHAEVALLSKOANG 660

RESULT 4
ID ANK1_MOUSE STANDARD; PRT; 1862 AA.
AC Q02357;
DT 01-NOV-1995 (Rel. 32, Created)
DE 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte;
RX MEDLINE=92345717; PubMed=1386265;
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT regulatory domain."
RL Mamm. Genome 3:281-285(1992).
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
```

```
CC -1- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M84756; AAA37236.1; -
CC HSSP; Q00420; IAWC.
CC MGD; MGI:88024; Ank1.
CC InterPro: IPR000488; -
CC InterPro: IPR000906; -
CC InterPro: IPR002110; -
CC Pfam; PF00791; ZUS; 1.
CC Pfam; PF00023; ank; 23.
CC Pfam; PF00531; death; 1.
CC PROSITE; PS50088; ANK_REPEAT; 20.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50017; DEATH_DOMAIN; 1.
CC Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
KW Cytoskeleton; Repeat; ANK repeat; ANION EXCHANGE PROTEIN
DOMAIN 1 827
FT FT 828 1386 62 KDA DOMAIN (SPECTRIN BINDING
FT FT 1387 1862 55 KDA REGULATORY DOMAIN (REGULATES
FT FT THE BINDING OF ANKYRIN TO SPECTRIN
FT FT AND THE BAND 3 PROTEIN).
FT FT ANK 1.
FT REPEAT 40 69 ANK 2.
FT REPEAT 73 102 ANK 3.
FT REPEAT 106 135 ANK 4.
FT REPEAT 139 168 ANK 5.
FT REPEAT 170 197 ANK 6.
FT REPEAT 201 230 ANK 7.
FT REPEAT 234 263 ANK 8.
FT REPEAT 267 296 ANK 9.
FT REPEAT 300 329 ANK 10.
FT REPEAT 333 362 ANK 11.
FT REPEAT 366 395 ANK 12.
FT REPEAT 399 428 ANK 13.
FT REPEAT 432 461 ANK 14.
FT REPEAT 465 494 ANK 15.
FT REPEAT 498 527 ANK 16.
FT REPEAT 531 560 ANK 17.
FT REPEAT 564 593 ANK 18.
FT REPEAT 597 626 ANK 19.
FT REPEAT 630 659 ANK 20.
FT REPEAT 663 692 ANK 21.
FT REPEAT 696 725 ANK 22.
FT REPEAT 729 758 ANK 23.
FT REPEAT 762 791 DEATH DOMAIN.
FT FT 1399 1483
FT SEQUENCE 1862 AA; 204242 MW; A56B85B5B29001E5 CRC64;

Query Match 24.7%; Score 287.5; DB 1; Length 1862;
Best Local Similarity 32.9%; Pred. No. 1.7e-17;
Matches 78; Conservative 31; Mismatches 89; Indels 39; Gaps 5;

QY 26 ILADK--SLATRTDQDSRTALHWACSAGTHIVEFLQLQGVNDKDGWS----- 75
Db 253 LLLDRGAQIETRT-KDELTPHCAARNGHVRISEILLDGHGAPIQAKTGLSLPIHMAAQG 311
QY 76 -----PLHIAASAGRDEIVKALLGKGAQVNAVNGCTPL 110
Db 312 DHLDCVRLQLQYNAEIDDTLDHLTPHVAACHGHHRVAKVLLDKGAKPNSRALNGFTPL 371
QY 111 HYAASKNRHEIAVMLLGGGANPDAKOHYEATAMHRAAKGNLKMHIILLYKASTNTQDT 170
```



```
Db 372 HIACKNHIRVMEILLKTGASIDAVTESGLATPLHVASPMGHLPIVKNLQRGASPNVSNV 431
Qy 171 EGNTPHLACDEEREEAKLLVSOGASIIYIENKEKTPLOVAK--GGGLGLILKRMVE 225
Db 432 KVETPLHMAARAGHTVAYKLLQNKAKANAKAKODQTPHCAARIGHTGMV-KLLLE 487

RESULT 5
YD57_SCHPO STANDARD; PRT; 234 AA.
AC Q10311;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 25.9 KDA PROTEIN C6C3.07 IN CHROMOSOME 1.
GN SPAC6C3.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69731; CA93620.1; -
CC HSSP; Q13625; 1YCS
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 5.
CC PROSITE; PS50088; ANK_REPEAT; 3.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC KW Hypothetical protein; ANK repeat; Repeat.
CC FT REPEAT 36 66 ANK 1.
CC FT REPEAT 70 100 ANK 2.
CC FT REPEAT 106 135 ANK 3.
CC FT REPEAT 140 169 ANK 4.
CC FT REPEAT 173 203 ANK 5.
CC SQ SEQUENCE 234 AA; 25930 MW; 828F6B00A0A0B0A08E CRC64;

Query Match 23.5%; Score 273.5; DB 1; Length 234;
Best Local Similarity 34.2%; Pred. No. 2.3e-17;
Matches 68; Conservative 35; Mismatches 87; Indels 9; Gaps 5;

Qy 20 BELKESILADKSLATRTDQDSRTALHWACSAGHTIVEFL-OLGVPVNDKDDAGWSPLH 78
Db 17 EIVEQAIQNDPSNLNAVDDKRTPLHWACSGKGVNTIYFLKQPNKIDKDEKAGWTPLM 76
Qy 79 IAAS--AGRDEIVKALLGK-AQVNAVQNGCTPLHYAASKNRHRIAYMLLEGGANPD-- 133
Db 77 TSINRSVPDNIIEINRSDDVPTITRGQTCLHYAAGKGRSLIVQLCDKA--PELI 134
Qy 134 -AKDHYETAMHRAAAKNLKMHIILYYKASTNIQDTGNTPLHACDEEREEAKLLV 192
Db 135 RKKDLQGTPLHRAAAGKIQVKKYLIISORAPLNTSDSYGFTPLHFAEGHPDVGVELV 194
Qy 193 SOGASIIYIENKEKTPLOV 211
Db 195 RAGATLTKRDSNHTALEV 213

RESULT 6
```

```
PA26_MOUSE
ID PA26_MOUSE STANDARD; PRT; 752 AA.
AC P97819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-PLA2).
GN PLA2G6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2;
RX MEDLINE=97236816; PubMed=9079688;
RA Balboa M.A., Balsinde J., Jones S.S., Dennis E.A.;
RT "Identity between the Ca2+-independent phospholipase A2 enzymes from
RT p38D1 macrophages and Chinese hamster ovary cells.";
RL J. Biol. Chem. 272:8576-8580(1997).
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERYL-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U8624; AAB48511.1; -
CC MGD; MGI:1859152; Pla2g6.
CC HSSP; P25963; 1NFI.
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 6.
CC PROSITE; PS50088; ANK_REPEAT; 4.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC KW Hydrolase; Lipid degradation; Repeat; ANK repeat.
CC FT REPEAT 151 181 ANK 1.
CC FT REPEAT 185 215 ANK 2.
CC FT REPEAT 219 248 ANK 3.
CC FT REPEAT 251 281 ANK 4.
CC FT REPEAT 286 312 ANK 5.
CC FT REPEAT 316 345 ANK 6.
CC FT REPEAT 349 378 ANK 7.
CC FT ACT_SITE 465 465 POTENTIAL.
CC SQ SEQUENCE 752 AA; 83728 MW; 7CBE9D574741478B CRC64;

Query Match 21.7%; Score 252.5; DB 1; Length 752;
Best Local Similarity 28.9%; Pred. No. 7.3e-15;
Matches 71; Conservative 40; Mismatches 94; Indels 41; Gaps 6;

Qy 2 EGCVENLWVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTIVEFLQ 61
Db 151 EGCFTPLHLACRGDSEILVELVQYCHAQMDV---TDNKGETAHYAVQDGNPOVLQGLK 207
Qy 62 LGVP-VNKKDDAGWPLHTAASAGRDEIVKALLGKAQVNAV-----NONGC 107
Db 208 NASPGELNVNNOGLTPLHLACKMGKQEMVRVULLLCNARCNINPGGFPHTAMKSKQCC 267
Qy 108 -----TPLHYAASKNRHRIAYMLLEGGANPDADKDYETAMHRA 146
Db 268 AEMIISMSNQIHSKDPYRGASPLHWA--KNA-ENARMLLKRGCDVDSTSSSGNTALHVA 324
Qy 147 AAKGNLKMHIILYYKASTNIQDTGNTPLHACDEEREEAKLLVSOGASIIYIENKEK 206
Db 325 VMNRNFDCCVMVLLTYGANAGARGEHNTPLHLAMSKDNMEMVKALIVFGAEDYDTPNDFGE 384
```

Best Local Similarity 29.6%; Pred. No. 2.1e-14;
Matches 64; Conservative 44; Mismatches 106; Indels 2; Gaps 2;

QY 207 TPLOVA 212
DB 385 TPALIA 390

RESULT 7
ANR3_HUMAN STANDARD; PRT; 832 AA.
AC P57078;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE ANKRD3 (EC 2.7.1.1-) (ANKYRIN REPEAT
DOMAIN PROTEIN 3).
GN ANKRD3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Sharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramer J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Vaspo M.-L.;
RL "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001743; BRA95526.1; -
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS00297; ANK_REPEAT; 1.
CC DR PROSITE; PS00088; ANK_REPEAT; 9.
CC KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
CC ANK repeat.
CC FT DOMAIN 22 286 PROTEIN KINASE.
CC FT REPEAT 485 514 ANK 1.
CC FT REPEAT 518 547 ANK 2.
CC FT REPEAT 551 580 ANK 3.
CC FT REPEAT 584 613 ANK 4.
CC FT REPEAT 617 647 ANK 5.
CC FT REPEAT 651 680 ANK 6.
CC FT REPEAT 684 713 ANK 7.
CC FT REPEAT 717 746 ANK 8.
CC FT REPEAT 750 780 ANK 9.
CC FT REPEAT 782 811 ANK 10.
CC FT NP_BIND 28 36 ATP (BY SIMILARITY).
CC FT BINDING 51 51 ATP (BY SIMILARITY).
CC FT ACT_SITE 143 143 BY SIMILARITY.
CC SQ SEQUENCE 832 AA; 91610 MW; 5D8FFED5F04F7ECB CRC64;
Query Match 21.3%; Score 248; DB 1; Length 832;

12 NLAYSGKLEELKESILADKSLATRTDQDSTALHWAACSAAGHTEIVFLQLQGVFVNDKDD 71
524 HMAVERRVRGVVELLLARKISVNAKDEDQWALHFAAQNGDESSTRLLLEKNASVNEVDF 583
72 AGWSPLHIAASAGRDEIVKALGKGAQVNAVNGCTPLHYAASKNRHEIAVMLL-EGGA 130
584 EGRTPMHVACQHQENIVRILLRRGVDSVLSQGDWALPLHYAAQOGLHPLTVKLLAKQPGV 643
131 NPAKDHYEATAMHRAAKGNLKMIIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKL 190
644 SVNAQTLDRGTPLHLAAQQRHYVARILIDLCSDVNVCSLLAQTPLHVAETGHTSTARL 703
191 LVSOQASIVIEKKEETPLQV-AKGGILGLTKRMVE 225
704 LLHRGAGKEAMTSDGYTALHLAARNHGLATVTKLIVE 739

RESULT 8
PA26_RAT STANDARD; PRT; 751 AA.
AC P97370;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-
PLA2).
GN PLA2G6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Pancreatic islets;
RX MEDLINE=97269008; PubMed=9111008;
RA Ma Z., Ramanadham S., Kempe K., Chi X.S., Ladenson J., Turk J.;
RT "Pancreatic islets express a Ca2+-independent phospholipase A2 enzyme
RT that contains a repeated structural homologous to the integral
RT membrane protein binding domain of ankyrin.";
RL J. Biol. Chem. 272:11118-11127(1997).
CC -1- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM PHOSPHOLIPIDS.
CC IT HAS BEEN IMPLICATED IN NORMAL PHOSPHOLIPID REMODELLING, NITRIC
CC OXIDE-INDUCED OR VASOPRESSIN-INDUCED ARACHIDONIC ACID RELEASE AND
CC IN LEUKOTRIENE AND PROSTAGLANDIN PRODUCTION. MAY PARTICIPATE IN
CC FAS MEDIATED APOPTOSIS AND IN REGULATING TRANSMEMBRANE ION FLUX IN
CC GLUCOSE-STIMULATED B-CELLS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: FOUND IN BRAIN, LUNG, SPLEEN, KIDNEY, LIVER,
CC HEART AND SKELETAL MUSCLE.
CC -1- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51898; AAC53136.1; -
CC HSSP; P42773; LIHB.
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 6.
CC PROSITE; PS50088; ANK_REPEAT; 4.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC KW Hydrolase; Lipid degradation; Repeat; ANK repeat.
CC REPEAT 150 180 ANK 1.
CC REPEAT 184 214 ANK 2.

```
FT REPEAT 218 247 ANK 3.
FT REPEAT 250 280 ANK 4.
FT REPEAT 285 311 ANK 5.
FT REPEAT 315 344 ANK 6.
FT REPEAT 348 377 ANK 7.
FT ACT_SITE 464 464 POTENTIAL.
SQ SEQUENCE 751 AA; 83582 MW; 393BBBADA7FCC99B CRC64;

Query Match 21.3%; Score 247.5; DB 1; Length 751;
Best Local Similarity 28.5%; Pred. No. 2e-14;
Matches 70; Conservative 39; Mismatches 96; Indels 41; Gaps 6;

QY 2 EGVSNLAVNCLAYSGKLEELKESILADKSLATRDQDSRTALHWACSGAGTEIVEEFL-L 60
DB 150 ECTPLHLACRGDSEILVELVOYCHAQMDV---TDNGETAFHYAVOGDNPVQLQGLK 206
QY 61 QLGVPNKDDAGNSPLHIAASAGDEIVKALLGKGAQVNAV-----NQNGC 107
DB 207 NASAGLNQVNNQGLTPLHLACQMGQEMVRVLLLCNARCNINGPGGFPPIHTAMKFSQKGC 266
QY 108 -----TPLHYASKNRHETAVMLLEGGANPDADKHYEATAMHRA 146
DB 267 AEMIIISMSNQIHSKDPRYGASPLHWA--KNA-EMARMLLKRGCDVDSTSGNTALHYA 323
QY 147 AAKGNLKMHIHLLYKASTNIODTEGNTPLHLACDEERVEEAKLLVSGASIIYIENKEEK 206
DB 324 VTRNRFCDVWLLTYGANAGARGEHNTPLHLAMSKDNEMVKALIVFGAEDVTPNDRGE 383
QY 207 TPLQVA 212
DB 384 TPAFIA 389

RESULT 9
LATA_LATMA
ID LATA_LATMA STANDARD; PRT; 1401 AA.
AC P23631;
DC STRAIN-TREDECIMGUTTATUS; TISSUE=Venom;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE ALPHA-LATROTOXIN PRECURSOR.
OS Latrodectus mactans (Black widow spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Theridiidae; Latrodectus.
OX NCBI_TaxID=6924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TREDECIMGUTTATUS; TISSUE=Venom;
RX MEDLINE=91031994; PubMed=1577615;
RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black widow spider venom."
RL FEBS Lett. 270:127-131(1990).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=91362695; PubMed=1888339;
RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT "Structure of tryptic fragments of a neurotoxin from black widow spider venom."
RL Bioorg. Khim. 17:437-441(1991).
CC -1- FUNCTION: CAUSES SYNAPTIC VESICLES EXOCYTOSIS AND NEUROTRANSMITTER RELEASE FROM PRESYNAPTIC NERVE TERMINALS.
CC -1- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.
CC -1- SIMILARITY: CONTAINS 21 ANK REPEATS.
DR PIR: S11527; S11527.
DR HSSP: Q00420; IAWC.
DR InterPro: IPR002110; .
DR Pfam: PF00023; ank; 17.
DR PROSITE: PS50088; ANK_REPEAT; 11.
DR PROSITE: PS50297; ANK_REPEAT; 1.
KW Toxin; Neurotoxin; Repeat; ANK repeat; Venom; Signal.
FT SIGNAL 1 20
```

```
FT CHAIN 21 1401 ALPHA-LATROTOXIN.
FT REPEAT 490 521 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 559 589 ANK 3.
FT REPEAT 593 622 ANK 4.
FT REPEAT 626 656 ANK 5.
FT REPEAT 660 690 ANK 6.
FT REPEAT 695 723 ANK 7.
FT REPEAT 729 758 ANK 8.
FT REPEAT 762 791 ANK 9.
FT REPEAT 795 824 ANK 10.
FT REPEAT 828 857 ANK 11.
FT REPEAT 862 891 ANK 12.
FT REPEAT 895 924 ANK 13.
FT REPEAT 928 957 ANK 14.
FT REPEAT 971 1003 ANK 15.
FT REPEAT 1003 1033 ANK 16.
FT REPEAT 1035 1064 ANK 17.
FT REPEAT 1068 1097 ANK 18.
FT REPEAT 1101 1131 ANK 19.
FT REPEAT 1137 1166 ANK 20.
FT REPEAT 1170 1199 ANK 21.
SQ SEQUENCE 1401 AA; 156834 MW; 21CE198B80172568 CRC64;

Query Match 21.3%; Score 247.5; DB 1; Length 1401;
Best Local Similarity 31.5%; Pred. No. 4.3e-14;
Matches 57; Conservative 34; Mismatches 87; Indels 3; Gaps 2;

QY 42 TALHWACSGAGTEIVEEFLQLGVPNVNDKDDAGNSPLHIAASAGDEIVKALLGKGAQVNA 101
DB 732 TPLHLAVIQGRKQILSLMFDIGVNIEQKTDKEYTPLHLAAMSKYPPELIQILLDQGSNFEA 791
QY 102 VNQNGCTPLHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLY 161
DB 792 KTNNGATPLHLATFKGSAALILLNNEVNRDITDENGQMPIHGAAMTGLLDVAQAIIIS 851
QY 162 KAS-TNIODTEGNTPLHLACDEERVEEAKLLVSGASIIYIENKEEKTPLQV--AKGGLGL 218
DB 852 DATVVDIEDKNSDTPNLNLAQNSHDIVIKYFIDQGDINTRNKKGLAPLLATSKGNLDM 911
QY 219 I 219
DB 912 V 912

RESULT 10
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
DC 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.-) (DAP KINASE 1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129831; PubMed=7828849;
RA Delass L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death."
RT death."
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL DEATH.
```

CC -1- PTM: AUTOPHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76104; CAA53712.1; -;
DR HSP; Q63450; 1A06.
DR MIM; 600831; -;
DR InterPro; IPR000488; -;
DR InterPro; IPR000719; -;
DR InterPro; IPR002110; -;
DR InterPro; IPR002290; -;
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00069; kinase; 1.
DR PROSITE; PS00088; ANK_REPEAT; 6.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
FT DOMAIN 13 266
FT PROTEIN_KINASE.
FT DOMAIN 267 334
FT REPEAT 378 407
FT REPEAT 411 440
FT REPEAT 444 473
FT REPEAT 478 507
FT REPEAT 511 540
FT REPEAT 544 573
FT REPEAT 577 606
FT REPEAT 610 639
FT REPEAT 639 876
FT REPEAT 876 905
FT REPEAT 1163 1197
FT DOMAIN 1313 1397
FT NP_BIND 19 27
FT BINDING 42 42
FT ACT_SITE 139 139
FT MUTAGEN 42 42
FT K->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1431 AA; 160017 MW; 9E84811004A15B CRC64;

Query Match 21.2%; Score 247; DB 1; Length 1431;
Best Local Similarity 29.0%; Pred. No. 4.9e-14;
Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;
Qy 37 DQDRTALHWACAGHTEIVFLQLQVGVNDKDGWSPHLHIAASGRD----- 87
Db 409 DGGSNVAVYWAARHGVDTLKFLSENKPLDVKDKSGEMALHVAARYGHADVAQVTCAS 468
Qy 88 -----IVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIA 122
Db 469 AQIPISRTKEETPLHCAWHGYYSVAKALCEAGCNVKNRGETPLLTASARGYHDIV 528
Qy 123 VMLGEGANPDADKHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTEGNTPLHLACDE 182
Db 529 ECLAEHGADLNACDKDGHIALHLAVRRCQMEVITKLLSQGCFVDYQDRHGNTPLHVACKD 588
Qy 183 ERVEAKLLVSGASIIYENKEEKTPLQVA 212
Db 589 GNMPIVVALCEANCLDINSYIGRTPLHLA 618
RESULT 11

AKR_ARATH
ID AKR_ARATH STANDARD; PRT; 439 AA.
AC Q05753;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN REPEAT PROTEIN (AKRP).
GN AKR.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RN STRAIN=CV. C24; TISSUE=Leaf;
RC MEDLINE=93104681; PubMed=1281700;
RA Zhang H., Scheirer D.C., Fowle W.H., Goodman H.M.;
RT 'Expression of antisense or sense RNA of an ankyrin repeat-containing
RL Plant Cell 4:1575-1588(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TEMPORAL AND SPATIAL REGULATION
CC OF CHLOROPLAST DEVELOPMENT FROM PROPLASTID.
CC -1- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION OCCURS IN TWO-WEEK-OLD
CC PLANTS AND DECLINES AS PLANTS DEVELOP FURTHER.
CC -1- SIMILARITY: CONTAINS 4 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M82883; AAA32812.1; -;
DR PIR; JQ1729; JQ1729.
DR HSP; Q13625; 1YCS.
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 4.
DR PROSITE; PS00088; ANK_REPEAT; 3.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
KW Cytoskeleton; Repeat; ANK repeat; Multigene family.
FT REPEAT 288 317
FT REPEAT 321 350
FT REPEAT 354 383
FT REPEAT 387 416
FT SEQUENCE 439 AA; 49150 MW; C371A90028B25BF3 CRC64;
Query Match 20.3%; Score 236.5; DB 1; Length 439;
Best Local Similarity 37.3%; Pred. No. 1e-13;
Matches 60; Conservative 19; Mismatches 73; Indels 9; Gaps 2;
Qy 57 EFL-----QLGVPVNDKDGWSPHLHIAASGRDEIVKALLGKGAQVNAVQNGCTPLH 111
Db 272 KFMLNSRNPDLAVATSK-----WLPLHTLAACGEFYLDVSLKHLNDINATDVGGLTALH 327
Qy 112 YAASKNRHEIATVLMLEGGANPDADKHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTE 171
Db 328 RAIIGKQATNYLLRSANFPVLDDEGATLMHYAVOTASAPTIKLLLYNADINAQDRD 387
Qy 172 GNTPLHLACDEERVEEAKLLVSGASIIYENKEEKTPLQVA 212
Db 388 GWTPLHVAVQARRSDIVKLLLIKAGDIEVKNKDGLTPLGEA 428
RESULT 12
ID PA26_HUMAN STANDARD; PRT; 806 AA.
AC Q06733; Q75645; Q9Y671; Q9UIT0; Q9UG29;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

Search completed: August 13, 2001, 07:44:50
Job time: 27 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 22.54 Seconds
(without alignments)
763.773 Million cell updates/sec

Title: US-09-509-775-2
Perfect score: 1164
Sequence: 1 MEGCVNLMVCNLAYSGKLE.....TPLQVAKGGLILKRWVEG 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	338	29.0	247	D84448	probable ankyrin [
2	312.5	26.8	1765	T42714	ankyrin 3, splice
3	312.5	26.8	1940	T42715	ankyrin 3, splice
4	312.5	26.8	1943	T42713	ankyrin 3, splice
5	312.5	26.8	1961	T42716	ankyrin 3, splice
6	310	26.6	237	T50984	related to 26s pro
7	310	26.6	4377	A55575	ankyrin 3, long sp
8	305	26.2	636	T33631	hypothetical prote
9	294	25.3	3924	S37431	ankyrin 2, neuro
10	291	25.0	228	S57697	hypothetical prote
11	290	24.9	1856	B35049	ankyrin 1, erythro
12	290	24.9	1880	A35049	ankyrin 1, erythro
13	290	24.9	1881	S3HUK	ankyrin 1, erythro
14	288.5	24.8	1786	A57282	ankyrin-related pr
15	288.5	24.8	1815	T15346	elegans ankyrin-re
16	288.5	24.8	1867	T15344	ankyrin-related un
17	288.5	24.8	2039	T15347	ankyrin-related un
18	287.5	24.7	1848	S37711	ankyrin, erythrocy
19	287.5	24.7	1862	T49502	ankyrin - mouse
20	276.5	23.8	1549	T13940	ankyrin - fruit fl
21	273.5	23.5	234	T39032	hypothetical ankyr
22	270.5	23.2	815	JG0197	myosin-light-chain
23	269.5	23.2	658	S58418	protein phosphatas
24	268.5	23.1	1004	A55142	myosin-light-chain
25	260.5	22.4	1062	T30255	inversin - mouse
26	248	21.3	397	T46445	hypothetical prote
27	248	21.3	1062	T14151	inv protein - mous
28	247.5	21.3	1401	S11527	alpha-latrotxin p
29	247	21.2	1423	I.137275	death-associated p

30	243.5	20.9	1411	2	S30355	alpha-latoinsecto
31	239.5	20.6	368	2	T18184	ankyrin repeat pro
32	239	20.5	319	2	A57291	cytokine inducible
33	239	20.5	557	2	T46507	hypothetical prote
34	237.5	20.4	791	2	T42691	hypothetical prote
35	236.5	20.3	439	2	JQ1729	ankyrin-repeat pro
36	233	20.0	934	1	H71274	probable ankyrin -
37	232	19.9	851	2	T12503	hypothetical prote
38	232	19.9	1398	2	T21884	hypothetical prote
39	230	19.8	1031	2	T43458	hypothetical prote
40	228	19.6	2352	2	T30201	Notch homolog prot
41	226.5	19.5	1001	2	S30385	G9a protein - huma
42	225.5	19.4	1188	2	T19552	hypothetical prote
43	225	19.3	323	2	B47169	ankyrin-like repea
44	224.5	19.3	900	2	A42024	transcription fact
45	224.5	19.3	933	2	S17233	transcription fact

ALIGNMENTS

RESULT 1

D84448
probable ankyrin [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84448
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84448
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: GB:AE002093; NID:q4335756; PIDN:AAD17433.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g03430
A:Map position: 2

Query Match 29.0%; Score 338; DB 2; Length 247;

Best Local Similarity 41.18; Pred. No. 6.1e-22;
Matches 90; Conservative 32; Mismatches 81; Indels 16; Gaps 6;

Qy 22 LKESILADKSLATRTDQDSRTALHWACSGHTEIVEFLL---QLGVPVNDKDDAGMSPLH 78

Db 31 LSEEQLS-KSLNFR-NEDGRSLHVAASFHGSQIVKLLSSSDSEAKTVINSKDEGMAPLH 88

Qy 79 IAAAGRDEIVKALGKGAOVNAVNGCTPLHYAASNHRHEIAVMVLLEGGANPAKHV 138

Db 89 SAAISGNAELVEVLLTRGADVNAKNNGRTALHYAASGRLEIAQLLLTHGAKINITKV 148

Qy 139 EATAMHRAAKGNKMIHLLYYKASTNIQDPEGNTPL--HLACDEERV-----EAKLL 191

Db 149 GCTPLHRAASVKLEVCBEFLIEEGAEIDATDKMGOTALMHSVICDDKQLKSMQVAFLL 208

Qy 192 VSQASIVIEKKEETPLQVAKGGLGLIL----KRWEG 226

Db 209 IRHGADVDEKGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 2

T42714
ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42714

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1765;

Best Local Similarity 35.5%; Pred. No. 1.3e-18;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAAFLDDHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGGCTPLHYAASKNRHETAVMLLEG 128

DB 577 AKSGSLTPLHVAAHYDNQKVALLLDQGSFHAAGKNGYTPHIAAKKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEETPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLHV 719

RESULT 3

T42715

ankyrin 3, splice form 3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42715

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

A:Reference number: Z22237; MUID:95340633

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1940 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 834/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1940;

Best Local Similarity 35.5%; Pred. No. 1.3e-18;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAAFLDDHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGGCTPLHYAASKNRHETAVMLLEG 128

DB 577 AKSGSLTPLHVAAHYDNQKVALLLDQGSFHAAGKNGYTPHIAAKKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEETPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLHV 719

RESULT 4

T42713

ankyrin 3, splice form 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

A:Reference number: Z22237; MUID:95340633

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 855/1

C:Function:

A:Note: major kidney ankyrin

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1943;

Best Local Similarity 35.5%; Pred. No. 1.3e-18;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAAFLDDHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGGCTPLHYAASKNRHETAVMLLEG 128

DB 577 AKSGSLTPLHVAAHYDNQKVALLLDQGSFHAAGKNGYTPHIAAKKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIHLLYYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEETPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLHV 719

RESULT 5

T42716

ankyrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42716

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

A:Reference number: Z22237; MUID:95340633

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 1-1961 <PET>
A;Cross-references: EMBL:L40632; NID:9710548; PID:9710552; PIDN:AAB01607.1
A;Experimental source: strain C57BL/6J; Kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
C;Superfamily: ankryrin; ankryrin repeat homology
C;Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1961;
Best Local Similarity 35.5%; Pred. No. 1.3e-19;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

Qy 42 TALHWACSGAGTEIVEFLQLQGVPNVDKDGASPLHIAASAGRDEIVKALLGK----- 95
Db 517 TPLHLAAREGHEDVNAFLLDHGASLSITTKGFTPLHVAAYKGLKLEVASILLQKSASDPA 576
Qy 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVMLEGG 128
Db 577 AGKSGLTPLHVAHYDNQKVALLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATSLLEY 636
Qy 129 GANPPAKDHYEATAMHRAAANGKLMHILLYKASTNIQDEGTPTPLHLACDSEERVEEA 188
Db 637 GADANAVTROGCIASVHLAAQEGHVDMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696
Qy 189 KLLVSOGASIVYENKEETPLQV 211
Db 697 EVLVNQGHAHVAQTKMGYTPLVH 719

RESULT 6
T50984
related to 26s proteasome subunit p28 [imported] - Neurospora crassa
N;Alternate names: protein B7F18.30
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C;Accession: T50984
R;Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A;Reference number: 225286
A;Accession: T50984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <SCH>
A;Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A;Experimental source: BAC clone B7F18; strain OR74A
C;Genetics:
A;Gene: NCSP:B7F18.30
A;Map position: 6
A;Introns: 17/1; 25/3; 68/3

Query Match 26.6%; Score 310; DB 2; Length 237;
Best Local Similarity 34.5%; Pred. No. 1.5e-19;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

Qy 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFL-QLQGVPNVDKDDA 72
Db 12 ARDGKASTVESLLNANPKLAQKDDGRPLHWCASVNRKEVVELLVNQGDPDVEDDM 71
Qy 73 GWSPLHIAASA-GRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 131
Db 72 GWTPEMISAKVSDAIIDLLSRGADINQTNHOSQATLHFIASKNIDLARKLLSPDKM 131
Qy 132 PD-----AKDHYEATAMHRAAANGKLMHILLYKASTNIQDEGTPTPLHLACDEERVE 186
Db 132 PKPASVRVKRKGQYPLHRAAAGSVPMINLLQHKSPINASDNAGYTPPLHVAEGHGH 191
Qy 187 EAKLLVSOGASI-----YIENKEETPLQV 211
Db 192 AAVALLKAGAEFTKMDGYLALDLPADPKVRRFIEKAEKEGIEL 237

RESULT 7
A55575
ankryrin 3, long splice form - human
N;Alternate names: ankryrin G
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankryrin-G. A new ankryrin gene with neural-specific isoforms localized at the
A;Reference number: A55575; MUID:95138209
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: GB:U13616; NID:9608024; PIDN:AAA64834.1; PID:9608025
C;Genetics:
A;Gene: GDB:ANK3
A;Cross-references: GDB:424503; OMIM:600465
A;Map position: 10q21-10q21
C;Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homol
C;Keywords: alternative splicing; peripheral membrane protein
F;73-105/Domain: ankryrin repeat homology <AN01>
F;106-138/Domain: ankryrin repeat homology <AN02>
F;139-171/Domain: ankryrin repeat homology <AN03>
F;172-200/Domain: ankryrin repeat homology <AN04>
F;201-233/Domain: ankryrin repeat homology <AN05>
F;234-266/Domain: ankryrin repeat homology <AN06>
F;267-299/Domain: ankryrin repeat homology <AN07>
F;300-332/Domain: ankryrin repeat homology <AN08>
F;333-365/Domain: ankryrin repeat homology <AN09>
F;366-398/Domain: ankryrin repeat homology <AN10>
F;399-431/Domain: ankryrin repeat homology <AN11>
F;432-464/Domain: ankryrin repeat homology <AN12>
F;465-497/Domain: ankryrin repeat homology <AN13>
F;498-530/Domain: ankryrin repeat homology <AN14>
F;531-563/Domain: ankryrin repeat homology <AN15>
F;564-596/Domain: ankryrin repeat homology <AN16>
F;597-629/Domain: ankryrin repeat homology <AN17>
F;630-662/Domain: ankryrin repeat homology <AN18>
F;663-695/Domain: ankryrin repeat homology <AN19>
F;696-728/Domain: ankryrin repeat homology <AN20>
F;729-761/Domain: ankryrin repeat homology <AN21>
F;762-794/Domain: ankryrin repeat homology <AN22>
F;795-827/Domain: ankryrin repeat homology <AN23>

Query Match 26.6%; Score 310; DB 2; Length 4377;
Best Local Similarity 37.0%; Pred. No. 5.7e-18;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 20 EELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLQGVPNVDKDDAGWSPLHI 79
Db 545 EDVAAFLLDHGASLSITTKGFTPLHVAAYKGLKLEVASILLQKSASPDAAKSGSLTPLHV 604
Qy 80 AASAGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 139
Db 605 AAHYDNQKVALLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATLLLEYGADANAVTRQG 664
Qy 140 ATAMHRAAANGKLMHILLYKASTNIQDEGTPTPLHLACDEERVEEAKLLVSOGASIY 199
Db 665 IASVHLAAQEGHVDMSVLLSRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGHVD 724
Qy 200 IENKEETPLQV 211
Db 725 AQTGMGYTPLVH 736

RESULT 8
T33631
hypotheical protein F40G9.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33631
R;Graves, T.; Sutterer, C.; Ozersky, P.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid F40G9.
A;Reference number: Z21378
A;Accession: T33631
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-636 <GRA>
A;Cross-references: EMBL:AF099919; PIDN:AAC68798.1; GSPDB:GN00021; CESP:F40G9.1
A;Experimental source: strain Bristol N2; clone F40G9
C;Genetics:
A;Gene: CESP:F40G9.1
A;Map position: 3
A;Introns: 21/3; 54/2; 90/3; 130/1; 157/3; 219/1; 260/3; 316/1; 351/3; 415/1; 438/3; 461/3

Query Match 26.2%; Score 305; DB 2; Length 636;
Best Local Similarity 29.3%; Pred. No. 1.4e-18;
Matches 86; Conservative 39; Mismatches 79; Indels 90; Gaps 6;

QY 5 VSNLWCVN-LAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLG 63
Db 335 ILHIKMKILHYFTKNVEARKLTRYPKLVGYTDDSGRSTIHEFAVGGSLPLLOFAI--- 391
QY 64 VPVNDKDDA----- 72
Db 392 --LNDPEMAHKTDVVRGMNLTKEIENLFENFKICDFRKNYSKNTKNPHSEFFETID 449
QY 73 -----GWSPLHIAASAGDEIVKALLG-KGAQVNA 101
Db 450 FEHNSONALKKFFSLKIDIFDLNCLILPLGWTPLMIASSAGRVVVRVYLLTLPDQVVKH 509
QY 102 VNONGCTPLHYAAKKNRHETAVMLLEGGAN-PDAKHYTEATAMHRAAAGNLMKMIHILLY 160
Db 510 TSNKQTSLSHYACSNHVEIVLKLLEADPNIIINLPKFGATLHRAASRGNDVIVRALVS 569
QY 161 Y-KASTNIQDTEGNTPLHLACDEERVEEAKLLVSOGASIYIENKEEKTPLQVAK 213
Db 570 TGKCSLDRODQEGNTALHLACDENRGDVAILLVNRGADMKMLNKEQTPLEMLK 623

RESULT 9
S37431
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C;Accession: S37431; A39643; B39643; A40334; A49462; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302466
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OTI>
A;Cross-references: GB:X56957
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443, 3585-3924 <OTT>
A;Cross-references: EMBL:X56958

R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Wa
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009921
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474, 'PE', 477-495 <TSE>
A;Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and
A;Reference number: A49462; MUID:94075409
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
C;Genetics:

A;Gene: GDB:ANK2
A;Cross-references: GDB:127607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F:2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F:63-95/Domain: ankyrin repeat homology <AN01>
F:96-128/Domain: ankyrin repeat homology <AN02>
F:129-161/Domain: ankyrin repeat homology <AN03>
F:162-190/Domain: ankyrin repeat homology <AN04>
F:191-223/Domain: ankyrin repeat homology <AN05>
F:232-267/Domain: ankyrin repeat homology <AN06>
F:265-297/Domain: ankyrin repeat homology <AN07>
F:298-330/Domain: ankyrin repeat homology <AN08>
F:331-363/Domain: ankyrin repeat homology <AN09>
F:364-396/Domain: ankyrin repeat homology <AN10>
F:397-429/Domain: ankyrin repeat homology <AN11>
F:430-462/Domain: ankyrin repeat homology <AN12>
F:463-495/Domain: ankyrin repeat homology <AN13>
F:496-528/Domain: ankyrin repeat homology <AN14>
F:529-561/Domain: ankyrin repeat homology <AN15>
F:562-594/Domain: ankyrin repeat homology <AN16>
F:595-627/Domain: ankyrin repeat homology <AN17>
F:628-660/Domain: ankyrin repeat homology <AN18>
F:661-693/Domain: ankyrin repeat homology <AN19>
F:694-726/Domain: ankyrin repeat homology <AN20>
F:727-759/Domain: ankyrin repeat homology <AN21>
F:760-792/Domain: ankyrin repeat homology <AN22>
F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 25.3%; Score 294; DB 2; Length 3924;
Best Local Similarity 31.9%; Pred. No. 1.2e-16;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

QY 14 AYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQ----- 61
Db 472 ARAGQVEVVR-CLLRNGALVDARAREEQTPHIAASAGHTEIVEFLLQMAHHPDAATTNG 530
QY 62 -----LQVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKAQVN 100
Db 531 YTPLHISAREGQVDVASVLLLEAGAAHSLATKGTTPHVAKYGSLDVAKLLQRRRAAD 590
QY 101 AVNONGCTPLHYAAKKNRHETAVMLLEGGANPDADKHYTEATAMHRAAAGNLMKMIHILLY 160
Db 591 SAGKNGLTPLHYAAHYDNQKVALLLLEKASPHATKNGYTPHIAAKKNQMQIASTLLN 650
QY 161 YKASTNIQDTEGNTPLHLACDEERVEEAKLLVSOGASIYIENKEEKTPLQVA 212
Db 651 YGAETNIVTKQGVTPHLLASQEGHTDMVTLLLDKGANIHMSTKSGLTSLHLA 702

RESULT 10
S57697

hypothetical protein YGR232w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G8564
C:Species: Saccharomyces cerevisiae
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Mar-2001
C:Accession: S57697; S64556; S63913
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 kb CRM1-YLM9-PET54-SMI1-PHO81-YHB4-PFK1 region
A:Reference number: S57680
A:Accession: S57697
A:Molecule type: DNA
A:Residues: 1-228 <VAN>
A:Cross-references: EMBL:X87941; NID:G886908; PIDN:CAA61182.1; PID:G886926
A:Experimental source: strain S288C
R:van der Aart, Q.J.M.; Steensma, H.Y.
submitted to the Protein Sequence Database, May 1996.
A:Reference number: S64541
A:Accession: S64556
A:Molecule type: DNA
A:Residues: 1-228 <VAV>
A:Cross-references: EMBL:Z73017; NID:g1323418; PIDN:CAA97260.1; PID:e243668; PID:g132342
A:Experimental source: Strain S288C
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
Yeast 12, 385-390, 1996
A:Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SMI1-PHO81-YHB4-PFK1 region
A:Reference number: S63896; MUID:96267763
A:Accession: S63913
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-228 <VAF>
A:Cross-references: EMBL:X87941; NID:G886908; PIDN:CAA61182.1; PID:G886926
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 7R
A:Note: YGR232w
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:71-103/Domain: ankyrin repeat homology <ANI>

Query Match 25.0%; Score 291; DB 2; Length 228;
Best Local Similarity 35.7%; Pred. No. 6.4e-18;
Matches 75; Conservative 36; Mismatches 91; Indels 8; Gaps 4;
Qy 5 VSNLMVCNLAYSGKLEELKESLADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGV 64
Db 1 MSNYPLOACHENEFKVOELHSPSLULQDQDGRPLHWSVSFQAHETISFLSKME 60
Qy 65 PVN---DRDDAGWSPLHIAASAGRDEIVKALGKGAQ--VNAVNGCGCTPLHYAASKNRH 119
Db 61 NVNLDYDPSDGSWTFPHIACSVGNLEVKSLVDRLKPLDKNITNQGVTCLEHLAVGKKWF 120
Qy 120 ETAVMLLEGGANPDADHYEATAMHRAAKGNLKMIIHLLYYKASTNIQDTTEGNTPLHL 178
Db 121 EVSQPLIENGASVRILKDFNQIPLHRAASVSGSLKIELLCGLGKSAVNQDQKQGTPLFH 180
Qy 179 ACDEERVEEAKLLVSQASIV--IENKEEK 206
Db 181 ALAEGHGDAVLLVEKYGAEDYLDVNDKGA 210

RESULT 11
B35049
ankyrin 1, erythrocyte splice form 3 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C:Accession: B35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: B35049

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1856 <LAM>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MA2>
F:2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 24.9%; Score 290; DB 2; Length 1856;
Best Local Similarity 34.3%; Pred. No. 1e-16;
Matches 79; Conservative 38; Mismatches 105; Indels 8; Gaps 3;
Qy 5 VSNLMV---CNLAYSGKLEELKESLADKSLATRTDQDSRTALHWACSGAGTEIVEFLQL 61
Db 432 VSNVAVETPLHMAARAGHTEVAKYLLQNKAKVNAKDDQTPHCAARTGHTNMVKLLLE 491
Qy 62 LGVPVNDKDDAGWSPLHIAASAGRDEIVKALGKGAQVNAVNGCGCTPLHYAASKNRHEI 121
Db 492 NNANPLATTAGTHTPLHIAAREGHVETVLALLEKASOACMTKGTPLHVAAYKGVKRV 551
Qy 122 AVMLLEGGANPDADHYEATAMHRAAKGNLKMIIHLLYYKASTNIQDTTEGNTPLHLACD 181
Db 552 AELLERDAHPNAACNGLTPLHVAHHNNLDIVKLLPRGSGSPHSPANNVYTPHIAAK 611
Qy 182 ERVEEAKLLVSQASIVYENKEETPLQVA--KGG---LGUILKRWVEG 226
Db 612 QNQVEVARSLQYGGSAESAESVQGVTPHLAAQAGHAEMVALLSKQANG 661

RESULT 12
A35049
ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>

A;Cross-references: GB:M28880

C;Genetics:

A;Gene: GDB:ANK1; ANK

A;Cross-references: GDB:118737; OMIM:182900

A;Map Position: 8p11.2-8p11.2

C;Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing; cytoskeleton

F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>

F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>

F;44-76/Domain: ankyrin repeat homology <AN01>

F;77-109/Domain: ankyrin repeat homology <AN02>

F;110-142/Domain: ankyrin repeat homology <AN03>

F;143-171/Domain: ankyrin repeat homology <AN04>

F;172-204/Domain: ankyrin repeat homology <AN05>

F;205-237/Domain: ankyrin repeat homology <AN06>

F;238-270/Domain: ankyrin repeat homology <AN07>

F;271-303/Domain: ankyrin repeat homology <AN08>

F;304-336/Domain: ankyrin repeat homology <AN09>

F;337-369/Domain: ankyrin repeat homology <AN10>

F;370-402/Domain: ankyrin repeat homology <AN11>

F;403-435/Domain: ankyrin repeat homology <AN12>

F;436-468/Domain: ankyrin repeat homology <AN13>

F;469-501/Domain: ankyrin repeat homology <AN14>

F;502-534/Domain: ankyrin repeat homology <AN15>

F;535-567/Domain: ankyrin repeat homology <AN16>

F;568-600/Domain: ankyrin repeat homology <AN17>

F;601-633/Domain: ankyrin repeat homology <AN18>

F;634-666/Domain: ankyrin repeat homology <AN19>

F;667-699/Domain: ankyrin repeat homology <AN20>

F;700-732/Domain: ankyrin repeat homology <AN21>

F;733-765/Domain: ankyrin repeat homology <AN22>

F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 24.9%; Score 290; DB 2; Length 1880;

Best Local Similarity 34.3%; Pred. No. 1.1e-16; Indels 8; Gaps 3;

Matches 79; Conservative 38; Mismatches 105; Indels 8; Gaps 3;

QY 5 VSNLMV---CNLAYSGKLEELKESILADKSLATRTDSDSTALHWACSAGHTEIVPELLQ 61

Db 432 VSNVKTETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPHCAARIGHTNNVKILLE 491

QY 62 LGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEI 121

Db 492 NNANPNLATTAGTTPHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPHCAARIGHTNNVKILLE 491

QY 122 AVMLLEGGANPDADKHVEATAMHRAAKGNLKMHIILLYKASTNIQDTGNTPLHLACD 181

Db 552 AELLERDAHPNAAGKNGLTPLHVAVHNNLDIVKLLPRGGSPHSPAWNNGYTPPLHIAAK 611

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGVTPPLHIAAQEGHAEWALLSKOANG 661

RESULT 13

SJHUK

N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R

N;Contains: ankyrin 2.2

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999

C;Accession: S08275; A33219; PC2220; A35443

R;Lux, S.E.; John, K.M.; Bennett, V.

Nature 344, 36-42, 1990

A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure

A;Reference number: S08275; MUID:90158830

A;Accession: S08275

A;Molecule type: mRNA

A;Residues: 1-1881 <LUI>

A;Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702

A;Accession: A33219

A;Molecule type: protein

A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,
'X',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LOX>

A;Note: 845-Arg and 1392-Thr were also found

R;Hermann, J.; Barel, M.; Frade, R.

Biochem. Biophys. Res. Commun. 204, 453-460, 1994

A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membr

A;Reference number: PC2220; MUID:95071348

A;Accession: PC2220

A;Molecule type: protein

A;Residues: 910-929 <HER>

R;Davis, L.H.; Bennett, V.

J. Biol. Chem. 265, 10589-10596, 1990

A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchang

A;Reference number: A35443; MUID:90285190

A;Accession: A35443

A;Molecule type: protein

A;Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-8

C;Genetics:

A;Gene: GDB:ANK1; ANK

A;Cross-references: GDB:118737; OMIM:182900

A;Map position: 8p11.2-8p11.2

C;Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing; phosphoprotein

F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>

F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>

F;2-827/Domain: 89K #status predicted <DOM1>

F;2-827/Region: anion exchange protein binding

F;44-76/Domain: ankyrin repeat homology <AN01>

F;77-109/Domain: ankyrin repeat homology <AN02>

F;110-142/Domain: ankyrin repeat homology <AN03>

F;143-171/Domain: ankyrin repeat homology <AN04>

F;172-204/Domain: ankyrin repeat homology <AN05>

F;205-237/Domain: ankyrin repeat homology <AN06>

F;238-270/Domain: ankyrin repeat homology <AN07>

F;271-303/Domain: ankyrin repeat homology <AN08>

F;304-336/Domain: ankyrin repeat homology <AN09>

F;337-369/Domain: ankyrin repeat homology <AN10>

F;370-402/Domain: ankyrin repeat homology <AN11>

F;403-435/Domain: ankyrin repeat homology <AN12>

F;436-468/Domain: ankyrin repeat homology <AN13>

F;469-501/Domain: ankyrin repeat homology <AN14>

F;502-534/Domain: ankyrin repeat homology <AN15>

F;535-567/Domain: ankyrin repeat homology <AN16>

F;568-600/Domain: ankyrin repeat homology <AN17>

F;601-633/Domain: ankyrin repeat homology <AN18>

F;634-666/Domain: ankyrin repeat homology <AN19>

F;667-699/Domain: ankyrin repeat homology <AN20>

F;700-732/Domain: ankyrin repeat homology <AN21>

F;733-765/Domain: ankyrin repeat homology <AN22>

F;766-798/Domain: ankyrin repeat homology <AN23>

F;828-1382/Region: 62K #status predicted <DOM2>

F;828-1382/Region: spectrin binding

F;1383-1881/Domain: 55K #status predicted <DOM3>

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 19.4 seconds
(without alignments)
239.867 Million cell updates/sec

Title: US-09-509-775-2
Perfect score: 1164
Sequence: 1 MEGCVSNLMVCNLAYSGLKLE.....TPLQVAKGGLILKRMVEG 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	310	26.6	1088	US-09-082-059-2	Sequence 2, Appli
2	294	25.3	1839	US-09-172-977-4	Sequence 4, Appli
3	287	24.7	352	US-09-065-474-139	Sequence 139, App
4	287	24.7	1745	US-09-031-485-33	Sequence 33, Appl
5	287	24.7	1745	US-08-847-429A-33	Sequence 33, Appl
6	287	24.7	1745	US-09-065-474-33	Sequence 33, Appl
7	286	24.6	843	US-09-172-977-3	Sequence 3, Appli
8	271	23.3	302	US-09-031-485-38	Sequence 38, Appl
9	271	23.3	302	US-08-847-429A-38	Sequence 38, Appl
10	271	23.3	302	US-09-065-474-38	Sequence 38, Appl
11	262	22.5	303	US-09-031-485-23	Sequence 23, Appl
12	262	22.5	303	US-08-847-429A-23	Sequence 23, Appl
13	262	22.5	303	US-09-065-474-23	Sequence 23, Appl
14	258.5	22.2	787	US-09-188-930-334	Sequence 334, App
15	247	21.2	1423	US-08-810-712-10	Sequence 10, Appl
16	242.5	20.8	752	US-08-281-193-2	Sequence 2, Appli
17	242.5	20.8	752	US-08-422-106-2	Sequence 2, Appli
18	242.5	20.8	752	US-08-733-716-2	Sequence 2, Appli
19	242.5	20.8	752	US-08-555-568B-2	Sequence 2, Appli
20	242.5	20.8	752	PCT-US95-08069-2	Sequence 2, Appli
21	241	20.7	348	US-09-031-485-28	Sequence 28, Appl
22	241	20.7	348	US-08-847-429A-28	Sequence 28, Appl
23	241	20.7	348	US-09-065-474-28	Sequence 28, Appl
24	237	20.4	387	US-08-484-575A-7	Sequence 7, Appli
25	237	20.4	387	US-08-477-459-7	Sequence 7, Appli
26	237	20.4	387	US-08-479-869-7	Sequence 7, Appli
27	237	20.4	387	US-08-486-414-7	Sequence 7, Appli

28	237	20.4	387	5	PCT-US94-01826A-7	Sequence 7, Appli
29	237	20.4	387	5	PCT-US94-02252A-7	Sequence 7, Appli
30	227.5	19.5	394	2	US-08-555-568B-17	Sequence 17, Appl
31	227.5	19.5	687	2	US-08-555-568B-21	Sequence 21, Appl
32	227.5	19.5	688	2	US-08-555-568B-23	Sequence 23, Appl
33	226.5	19.5	191	2	US-09-031-485-20	Sequence 20, Appl
34	226.5	19.5	191	2	US-08-847-429A-20	Sequence 20, Appl
35	226.5	19.5	191	3	US-09-065-474-20	Sequence 20, Appl
36	224.5	19.3	899	1	US-08-365-689-2	Sequence 2, Appli
37	224.5	19.3	899	1	US-08-145-138A-2	Sequence 2, Appli
38	224.5	19.3	933	1	US-07-747-781-2	Sequence 2, Appli
39	224.5	19.3	933	5	PCT-US92-06888-2	Sequence 2, Appli
40	212.5	18.3	657	1	US-08-264-534-34	Sequence 34, Appl
41	212.5	18.3	657	1	US-08-083-590A-13	Sequence 13, Appl
42	212.5	18.3	657	1	US-08-465-500-34	Sequence 34, Appl
43	212.5	18.3	657	2	US-08-346-128-34	Sequence 34, Appl
44	212.5	18.3	657	3	US-08-532-384-13	Sequence 13, Appl
45	212.5	18.3	657	3	US-08-893-828-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identifi
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 26.6%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 1.9e-26;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;
QY 20 RELKESILADKSLATRTDQDSRTALHWACSGHTEIVEFLQLGLGVPNKDDAGWSPLHI 79
Db 166 EDVAAFLLDHGASLSITTKKGTPLHVAKYCKLEVANLLLOKSASPDANGSGLTPLHV 225
QY 80 AASAGRDEITVALLKGAGQAVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDARDHYE 139
Db 226 AAHYDNQKVALLLDQASPHAAKNGYTPHIAAKNQMDIATTLLEYGADANAVTROG 285
QY 140 ATAMHRAAKGNKMIHILLYKASTNTQDTGCTNPLHLACDEERVEEAKLLVSGASTY 199
Db 286 TASVHLAAQEGHVMVSLLLGRNANVNLNSKLSGLTPLHLAAQEDRYNVAEVLVNOGAHYD 345
QY 200 IENKEETPLQV 211
Db 346 AQTKMGYTPLVH 357

RESULT 2
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.

```
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
; US-09-172-977-4

Query Match      25.3%; Score 294; DB 2; Length 1839;
Best Local Similarity 31.9%; Pred. No. 2.9e-24;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

QY 14 AYSGLLEELKESILAKSLATRTDQSRTHLWACSGAGHTEIVEFLQ----- 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 472 ARAGQVEVVR-CLLRNGALVDARAREEQTPLHTASRLGKTEIVQLLQHHMAHPDAATNG 530
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 -----LGPVNDKDDAGWSPLHTAASAGRDEIVKALLGKGAQVN 100
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 531 YTPLHTSAREGVQDVASVLLLEAGAAHSLATKGTPLHVAAKYGSIDVAKLLQRRAAD 590
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 AVNQNGCTPLHYAAKSNRHEIAYMLLEGGANPDADKHYEATAMHRAAKGNLKMHIILY 160
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 591 SACKNGLTPLHVAHYNDQKVALLLLEKGAHPHATAKNGVTPHIAAKNQMDIASTLLN 650
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 161 YKASTNIQDTGNTPLHLACDEERVEEAKLLVSGASIVYENKEETPLQVA 212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 651 YGAETNIVTKOGVTPPLHLASQEGHDMVTLLLDKGANIHMSTRKGLTSLHLA 702
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-09-065-474-139
; Sequence 139, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-065-474-139

Query Match      24.7%; Score 287; DB 3; Length 352;
Best Local Similarity 36.5%; Pred. No. 1.4e-24;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 20 BELKESILA-----DKSLATRTDQSRTHLWACSGAGHTEIVEFLQLGVPVNDKDDA 72
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 KEGQEVAAALMDHGTDTLLTK---KGFTPLHAAKYGNLPAKSLLEGRGTPVDIEGKN 142
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 GWSPLHTAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAYMLLEGGANP 132
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 QVTPLHVAAYHNNDKVALLLENGASAAHAAKNGVTPHIAAKNQMDIASTLLHYKANA 202
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 DAKDHYEATAMHRAAKGNLKMHIILLYKASTNIQDTGNTPLHLACDEERVEEAKLLV 192
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 203 NAESKAGFTPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCQAQEDRVSVAEELV 262
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 SOGASIVYENKEETPLQVA 212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 KENNAIDPKTKAGYPLHVA 282
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

Qy 73 GWSPLHTAASAGRDEIVKALLGKAQVNAVQNCGCTPLHYAAASKNRHETAVMLLEGGANP 132
 ::: ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Dd 578 QVTPLHVAAAHYNDKVALLLLENCSAHAAKNGYTPHLIAAKKNQMIDIASTLLHYKAN A 637

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/847.429A
;; FILING DATE: 24-APR-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Verser, Carol Talkington
;; REGISTRATION NUMBER: 37,459
;; REFERENCE/DOCKET NUMBER: HW-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-847-429A-38

Query Match 23.3%; Score 271; DB 2; Length 302;
Best Local Similarity 35.4%; Pred. No. 7.4e-23;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 24 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPLHIAASA 83
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQOGANANVATVRGETPLHLAARA 116
Qy 84 GRDEIVKALLGGAQVNAVNGCTPLHYAASKNRHETAVMLLEGANPDADKHVEATAM 143
Db 117 NOTDIVRVLRNGAQQVDAARELQTPPLHIASRLGNTDIVILLQADASPNAATRDLYTL 176
Qy 144 HRAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIYIENK 203
Db 177 HIAAKEGEEVAAIILDHGSKTLTKGFTPLHLAAKYGNLPVAKLLERGTGLVDIEGK 236
Qy 204 EEKTPLOVA 212
Db 237 NQVTPPLHVA 245

RESULT 10
US-09-065-474-38
; Sequence 38, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065.474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-065-474-38

Query Match 23.3%; Score 271; DB 3; Length 302;
Best Local Similarity 35.4%; Pred. No. 7.4e-23;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 24 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPLHIAASA 83
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQOGANANVATVRGETPLHLAARA 116
Qy 84 GRDEIVKALLGGAQVNAVNGCTPLHYAASKNRHETAVMLLEGANPDADKHVEATAM 143
Db 117 NOTDIVRVLRNGAQQVDAARELQTPPLHIASRLGNTDIVILLQADASPNAATRDLYTL 176
Qy 144 HRAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIYIENK 203
Db 177 HIAAKEGEEVAAIILDHGSKTLTKGFTPLHLAAKYGNLPVAKLLERGTGLVDIEGK 236
Qy 204 EEKTPLOVA 212
Db 237 NQVTPPLHVA 245

RESULT 11
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-23

Query Match 22.5%; Score 262; DB 2; Length 303;
Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;
QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLGVPVND 68
DQ 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128
DQ 70 TTESGLSPLHVAAPMGAINIVYLLQQGANADVATVGTETPLHIAARANQTDIVRVLRN 129
QY 129 GANPDADKHDEYATAMHRAAAKGNLKMIIHLLYKASTNIQDTEGNTPLHLACDEERVEEA 188
DQ 130 GAQVDAARAELQPLHIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAAKGQEEVA 189
QY 189 KLLYSQAGSIYIENKEKTPLOV-AKGLGLILKRWVE 225
DQ 190 AILMDHGTDKTLTKGFTPLHIAAKYGNLPVAKSLLE 227

RESULT 12
US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-23

Query Match 22.5%; Score 262; DB 2; Length 303;

Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;
QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLGVPVND 68
DQ 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128
DQ 70 TTESGLSPLHVAAPMGAINIVYLLQQGANADVATVGTETPLHIAARANQTDIVRVLRN 129
QY 129 GANPDADKHDEYATAMHRAAAKGNLKMIIHLLYKASTNIQDTEGNTPLHLACDEERVEEA 188
DQ 130 GAQVDAARAELQPLHIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAAKGQEEVA 189
QY 189 KLLYSQAGSIYIENKEKTPLOV-AKGLGLILKRWVE 225
DQ 190 AILMDHGTDKTLTKGFTPLHIAAKYGNLPVAKSLLE 227

RESULT 13
US-09-065-474-23
; Sequence 23, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-CI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-23

Query Match 22.5%; Score 262; DB 3; Length 303;
Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;

QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLGVPVND 68
DQ 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:24 ; Search time 32.54 Seconds
(without alignments)
918.898 Million cell updates/sec

Title: US-09-509-775-2
Perfect score: 1164
Sequence: 1 MEGCVSNLMVCNLAYSGKLE.....TPLQVAKGGIGLILKRWVEG 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL16.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-unclassified.*
- 13: sp-vertebrate.*
- 14: sp-virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1164	100.0	226	4	075832 homo sapien
2	1106	95.0	231	11	Q922X3 rattus norv
3	1105	94.9	231	11	Q922X2 mus musculu
4	582.5	50.0	118	4	095533 homo sapien
5	338	29.0	247	10	Q92Q79 arabidopsis
6	312.5	26.8	1943	11	Q61307 mus musculu
7	310	26.6	237	3	Q9P3P6 neurospora
8	310	26.6	1088	4	Q13484 homo sapien
9	310	26.6	4377	4	Q12955 homo sapien
10	306	26.3	2622	11	Q70511 rattus norv
11	305	26.2	636	5	Q9T267 caenorhabdi
12	301	25.9	1762	11	Q88521 rattus norv
13	297	25.5	1159	5	Q9NCP8 drosophila
14	297	25.5	2443	5	Q9VSA2 drosophila
15	290.5	25.0	1136	6	Q9N180 bos taurus
16	290	24.9	1719	4	Q13768 homo sapien
17	290	24.9	1856	4	Q99407 homo sapien
18	288.5	24.8	1786	5	Q17344 caenorhabdi
19	288.5	24.8	1809	5	Q17487 caenorhabdi

20	288.5	24.8	1815	5	Q17488 caenorhabdi
21	288.5	24.8	1867	5	Q17486 caenorhabdi
22	288.5	24.8	2039	5	Q17489 caenorhabdi
23	288.5	24.8	6994	5	Q17343 caenorhabdi
24	287.5	24.7	1848	11	Q61302 mus musculu
25	286	24.6	843	11	P97582 rattus norv
26	285.5	24.5	1098	11	Q61304 mus musculu
27	283.5	24.4	1059	4	Q15084 homo sapien
28	283.5	24.4	1166	4	Q9H2K2 homo sapien
29	283.5	24.4	1265	4	Q9HAS4 homo sapien
30	281.5	24.2	1327	4	Q95271 homo sapien
31	281	24.1	1181	5	Q9XZ37 drosophila
32	281	24.1	1181	5	Q9VBP3 drosophila
33	276.5	23.8	1549	5	Q24241 drosophila
34	276.5	23.8	1549	5	Q9V4B1 drosophila
35	270.5	23.2	815	6	Q9TV77 sus scrofa
36	270.5	23.2	1030	4	Q14974 homo sapien
37	269.5	23.2	658	11	Q62937 rattus norv
38	269.5	23.2	976	11	Q10728 rattus norv
39	268.5	23.1	963	13	Q90624 gallus gall
40	268.5	23.1	1004	13	Q90623 gallus gall
41	266.5	22.9	1395	5	Q9XZC0 latrodectus
42	266	22.9	768	4	Q9ULJ7 homo sapien
43	263.5	22.6	2119	5	Q9VAU5 drosophila
44	260.5	22.4	1062	11	Q89019 mus musculu
45	260	22.3	1430	11	Q9JJP7 mus musculu

ALIGNMENTS

RESULT 1
075832 PRELIMINARY; PRT; 226 AA.
ID 075832;
AC 075832;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 26S PROTEASOME SUBUNIT P28.
GN DJ889N15.2 OR HUMAN GANKYRIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98382579; PubMed=9714768;
RA Hori T., Kato S., Saeki M., Demartino G.N., Slaughter C.A.,
RA Takeuchi C., Toh E.A., Tanaka K.;
RT "cDNA cloning and functional analysis of p28 (Nas6p) and p40.5
RT (Nas7p), two novel regulatory subunits of the 26S proteasome.";
RL Gene 216:113-122(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Enhanced expression of a novel tumour marker in the human
RT hepatomas.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB009619; BAA33215.1;
DR EMBL; AL031177; CAA20117.1;
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome.

```
SQ SEQUENCE 226 AA; 24428 MW; 57158E33146EC7C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1164; DB 4; Length 226;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
DB 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
QY 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVEG 226

RESULT 2
ID Q922X3 PRELIMINARY; PRT; 231 AA.
AC Q922X3;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DE GANKYRIN HOMOLOGUE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022014; BAA36954.1; -
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS00088; ANK_REPEAT; 5.
DR PROSITE; PS0297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;

Query Match
Best Local Similarity 95.08; Score 1106; DB 11; Length 231;
Matches 213; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
DB 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
QY 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225
```

```
RESULT 3
Q922X2
ID Q922X2 PRELIMINARY; PRT; 231 AA.
AC Q922X2;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
DE GANKYRIN.
GN PSMD10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of mouse gankyrin containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022022; BAA36969.1; -
DR HSSP; P42773; 1IHB.
DR MGD; MGI:1858898; Psmd10.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS00088; ANK_REPEAT; 5.
DR PROSITE; PS0297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 25115 MW; 5A3DB7027B0694E6 CRC64;

Query Match
Best Local Similarity 94.9%; Score 1105; DB 11; Length 231;
Matches 210; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
DB 1 MEGCVSNLWCVNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLL 60
QY 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225

RESULT 4
Q95533
ID Q95533 PRELIMINARY; PRT; 118 AA.
AC Q95533;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
DE DJ889N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
GN {PUTATIVE PARTIAL ISOFORM 2} (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031177; CAA20118.1; -
DR HSSP; Q13625; 1VCS.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 2.
```

```

Qy 192 VSGASITYENKEEKPLOVAGGLGLIL-----KRWEG 226
      :  || : || || | | | | | | | | | | | | | |
Db 209 IRHGADVDEKGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 6
ID Q61307 PRELIMINARY; PRT: 1943 AA.
AC Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; Q08866; Q08867;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ANKYRIN 3 (ANKYRIN G) (EPITHELIAL ANKYRIN) (ANKYRIN-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=95340633; PubMed=915634;
RA Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Vialamas M.,
RT Turtzo L.C., Otsuka A.J., Lux S.E.;
RA "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT ankyrin gene family and the major ankyrin in kidney, is expressed in
RT alternatively spliced forms, including forms that lack the repeat
RT domain.";
RL J. Cell Biol. 130:313-330(1995).
[2]
RN RP SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RC STRAIN=C3H/HEJ; TISSUE=BONE MARROW;
RX MEDLINE=97213781; PubMed=9060470;
RA Hook T.C., Peters L.L., Lux S.E.;
RT "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate
RT with mouse macrophage lysosomes.";
RL J. Cell Biol. 136:1059-1070(1997).
CC -1- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS: 1 (SHOWN HERE), 2, 3,
CC 4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC -1- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR EMBL; L40631; AAB01603.1; -
DR EMBL; L40632; AAB01605.1; -
DR EMBL; U89275; AAB58381.1; -
DR EMBL; L40631; AAB01602.1; -
DR EMBL; L40632; AAB01604.1; -
DR EMBL; L40632; AAB01606.1; -
DR EMBL; L40632; AAB01607.1; -
DR EMBL; U89274; AAB58380.1; -
DR HSSP; P55273; 1BD8.
DR MGD; MGI:88026; Ank3.
DR InterPro; IPR000488; -
DR InterPro; IPR000906; -
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00005; DEATH; 1.

```

KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Multigene family.
KW DOMAIN 1 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT DOMAIN 23 810 24 X ANK MOTIF REPEATS.
FT DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.
FT DOMAIN 1438 1943 REGULATORY DOMAIN (REGULATES BINDING OF
ANKYRIN TO SPECTRIN AND THE BAND 3
PROTEIN).
FT REPEAT 23 55 ANK MOTIF 1.
FT REPEAT 56 88 ANK MOTIF 2.
FT REPEAT 89 121 ANK MOTIF 3.
FT REPEAT 122 154 ANK MOTIF 4.
FT REPEAT 155 183 ANK MOTIF 5.
FT REPEAT 184 216 ANK MOTIF 6.
FT REPEAT 217 249 ANK MOTIF 7.
FT REPEAT 250 282 ANK MOTIF 8.
FT REPEAT 283 315 ANK MOTIF 9.
FT REPEAT 316 348 ANK MOTIF 10.
FT REPEAT 349 381 ANK MOTIF 11.
FT REPEAT 382 414 ANK MOTIF 12.
FT REPEAT 415 447 ANK MOTIF 13.
FT REPEAT 448 480 ANK MOTIF 14.
FT REPEAT 481 513 ANK MOTIF 15.
FT REPEAT 514 546 ANK MOTIF 16.
FT REPEAT 547 579 ANK MOTIF 17.
FT REPEAT 580 612 ANK MOTIF 18.
FT REPEAT 613 645 ANK MOTIF 19.
FT REPEAT 646 678 ANK MOTIF 20.
FT REPEAT 679 711 ANK MOTIF 21.
FT REPEAT 712 744 ANK MOTIF 22.
FT REPEAT 745 777 ANK MOTIF 23.
FT REPEAT 778 810 ANK MOTIF 24.
FT VARSPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 834 855 LVRK (IN ISOFORM 5).
FT VARSPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND
ISOFORM 3).
FT VARSPLIC 852 852 G -> GDKCTWFKIPKQVEVLVKS (IN ISOFORM 4
AND ISOFORM 6).
FT VARSPLIC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND
ISOFORM 7).
SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

Query Match 26.8%; Score 312.5; DB 11; Length 1943;
Best Local Similarity 35.5%; Pred. No. 5e-19; Mismatches 67; Indels 33; Gaps 1;
Matches 72; Conservative 31;

QY 42 TALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPHLTAASAGRDEIVKALLGK----- 95
DB 517 TPLHLAAREGHEDVAFLLDHGLSITTKGFTPLHVAAYKGLVASLLQLKSASPDA 576
QY 96 -----GAQVNAVNGNCTPLHYAASKNRHEIAVMLEGG 128
DB 577 AGKSLTPLHVAHYDNDKVALLLLDQAGSPHAAKNGYTPHLHIAAKNQMDIATSLLEY 636
QY 129 GANPKADHYEATAMHRAAAGNLKMIHLLYYKASTNIQDTEGNTPLHLACDEERVEE 188
DB 637 GADANAVTQGIASVHLAAGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSQASIIYENKEETPLQV 211
DB 697 EVLVNQAQHVDAQTGMGTPLHV 719

RESULT 7
QY9P3P6 PRELIMINARY; PRT; 237 AA.
AC Q9P3P6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RELATED TO 26S PROTEASOME SUBUNIT P28.
GN B7F18.30.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389891; CAB97304.1; -;
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 6.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome.
SQ SEQUENCE 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;

Query Match 26.6%; Score 310; DB 3; Length 237;
Best Local Similarity 34.5%; Pred. No. 4.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

QY 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL-QLGVPVNDKDDA 72
DB 12 ARDGKASIVESLLNANPKLAQRKDDGRLPLTHWACSYNRKEVELLVNQKGFDPDVEDDM 71
QY 73 GWSPLHTAASA-GDEIVKALLGKGAOVNAVNGNCTPLHYAASKNRHEIAVMLEGGAN 131
DB 72 GWTFFMISASVKSDSAIDLLSLSGADINOTNHSQFALHFIASKNIDILARKLLSPDMK 131
QY 132 PD-----AKDHYEATAMHRAAAGNLKMIHLLYYKASTNIQDTEGNTPLHLACDEERVE 186
DB 132 PKPASVRVKRGQVPLHRAAAGISGVPMLNLLQHKSPINASDNAGVTPHLHVAEGHGH 191
QY 187 EAKLLVSQASII-----YIENKEETPLQV 211
DB 192 AAVALLKAGAEETDKKMDGYLALDLPADPKVRRFIEKEAKEGIEL 237

RESULT 8
QY9P3P6 PRELIMINARY; PRT; 1088 AA.
AC QY9P3P6;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ANKYRIN G119.
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=96234670; Pubmed=8666667;
RA Kashgarian M., Morrow J.S., Ardito T., Stabach P.R., Mann A.S.,
RA Devatajan P.;
RT "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
and muscle that binds beta 1 sigma spectrin and associates with the
Golgi apparatus.";
RL J. Cell Biol. 133:819-830(1996).
DR EMBL; U43965; AAB08437.1; -;
DR HSP; P55273; IBD8.
DR InterPro; IPR000906; -;

```
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 12.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS00088; ANK_REPEAT; 12.
DR PROSITE: PS0297; ANK_REP_REGION; 1.
DR SMART: SM00218; ZU5; 1.
SQ SEQUENCE 1088 AA; 119427 MW; 769C88D40A78DE86 CRC64;

Query Match      26.6%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 3.8e-19;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 20 BELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWSPLHI 79
Db 166 EDVAFLDHGASUSITTKGFTPLHVAAYKYGKLEVANLLQKSAPDAAGKSGLTPLHV 225

Qy 80 AASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGGANPDADKHYE 139
Db 226 AAHYDNQKVALLLDQASPHAAAKNGYTPHIAAKKNQMDIATTTLLEYGADANAVTRQG 285

Qy 140 ATAMHRAAAGNLMKIMHILLYKASTNTQDTGNTPLHLACDEERVEAKLLVSGASIY 199
Db 286 IASVHLAAQEGHVDVMSLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 345

Qy 200 IENKEEKTPLQV 211
Db 346 AQTRMGYTPLHV 357

RESULT 9
Q12955 PRELIMINARY; PRT: 4377 AA.
AC Q12955;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ANKYRIN G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SKELETAL MUSCLE;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RL J. Cell Sci. 0:0-0(1998).
DR EMBL: AF102552; AAC78143.1; -.
DR EMBL: AF065150; AAC18853.1; -.
DR HSP; P55273; 1BD8.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS0297; ANK_REP_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZU5; 1.
FT NON_TER 2622 2622
SQ SEQUENCE 4377 AA; 480399 MW; F42379E5768B684 CRC64;

Query Match      26.6%; Score 310; DB 4; Length 4377;
Best Local Similarity 37.0%; Pred. No. 2.5e-18;
Matches 71; Conservative 36; Mismatches 86; Indels 0; Gaps 0;
```

```
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 20 BELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWSPLHI 79
Db 545 EDVAFLDHGASUSITTKGFTPLHVAAYKYGKLEVANLLQKSAPDAAGKSGLTPLHV 604

Qy 80 AASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGGANPDADKHYE 139
Db 605 AAHYDNQKVALLLDQASPHAAAKNGYTPHIAAKKNQMDIATTTLLEYGADANAVTRQG 664

Qy 140 ATAMHRAAAGNLMKIMHILLYKASTNTQDTGNTPLHLACDEERVEAKLLVSGASIY 199
Db 665 IASVHLAAQEGHVDVMSLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 724

Qy 200 IENKEEKTPLQV 211
Db 725 AQTRMGYTPLHV 736

RESULT 10
Q70511 PRELIMINARY; PRT: 2622 AA.
AC Q70511;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1999 (Tremblrel. 11, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE 270 KDA ANKYRIN G ISOFORM (ANKYRING) (FRAGMENT).
GN ANK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98417633; PubMed=9744885;
RA Zhang X., Bennett V.;
RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
RT requires multiple ankyrin G-specific domains.";
RL J. Cell Biol. 142:1571-1581(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Carpenter S.S., Zhang X.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SKELETAL MUSCLE;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RL J. Cell Sci. 0:0-0(1998).
DR EMBL: AF102552; AAC78143.1; -.
DR EMBL: AF065150; AAC18853.1; -.
DR HSP; P55273; 1BD8.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS0297; ANK_REP_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
DR SMART: SM00248; ANK; 1.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZU5; 1.
FT NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

Query Match      26.3%; Score 306; DB 11; Length 2622;
Best Local Similarity 36.5%; Pred. No. 2.9e-18;
Matches 70; Conservative 36; Mismatches 86; Indels 0; Gaps 0;
```

```
Qy 20 BELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWSPLHI 79
Db 166 EDVAFLDHGASUSITTKGFTPLHVAAYKYGKLEVANLLQKSAPDAAGKSGLTPLHV 225
```

```
Db 553 EDVAAFLLDHGASLITTKGFTPLHVAAYKGLKLEVASLLQKSPDAAGKSGLTPLHV 612
QY 80 AASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDADKHYE 139
Db 613 AAHYDNQKVALLLDQOGASPHAAKNGYTPPLHAAKKNQMDIATSLLEYGADANPYTRQG 672
QY 140 ATAMHRAAAKGNLKMTHILYYKASTNIQDTEGTPHLACDEERVEEAKLLYSOGASIY 199
Db 673 IASVHLAAQEGHVDMSLLSRNANVLSNKSGLTPLHLAAQEDRVNAEVLNQGAHVD 732
QY 200 IENKEEKTPLQV 211
Db 733 AQTMGYTPPLHV 744

RESULT 11
Q9T267 PRELIMINARY; PRT: 636 AA.
AC Q9T267;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE F40G9.1 PROTEIN.
GN F40G9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dean S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kersey J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Shaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RN Nature 368:32-38(1994).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Graves T., Sutterer C., Ozersky P.;
RT "The sequence of C. elegans cosmid F40G9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099919; AAC68798.1; -.
DR HSSP; Q00421; LAWG.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 636 AA; 73229 MW; AB55162AF5D5B1C0 CRC64;

Query Match 26.2%; Score 305; DB 5; Length 636;
Best Local Similarity 29.3%; Pred. No. 5.2e-19;
Matches 86; Conservative 39; Mismatches 79; Indels 90; Gaps 6;

QY 5 VSNLMVCN-LAYSGLKEELKESILADSLATRTDQDSRTALHWACSGHTEIVEIFLQLG 63
Db 335 ILHKMKCILHYFTKNVYAKRLLTRYKPLVGYTDDSGRSTIHFANVGSPLPLQFAI--- 391
```

```
QY 64 VPVNDKDDA----- 72
Db 392 --LNDPEMAHKTDVVRVGMNLTKEIENLFENRKICDFRKNYSKNTKNPHSEFFETID 449
QY 73 -----GWSPLHIAASAGRDEIVKALLG-KGAOVNA 101
Db 450 FEHENSQNAKKFFSLKIDIFDLNCLILPLGWTPLMIASSAGRVVRYLLTLPDQDVKH 509
QY 102 VNONGCTPLHYAASKNRHEIAVMLLEGAN-PDAKHYEATAMHRAAAKGNLKMTHILY 160
Db 510 TNSNKOTSILHYACSKNHVEIVKLLIEADPNINLPKFGATALHRAASRGNDVIVRALVS 569
QY 161 Y-KASTNIQDTEGTPHLACDEERVEEAKLLYSOGASIYIENKEEKTPLQVAK 213
Db 570 TGKCSLDRDQGEGETALHLCADENRGDVAILLVNRGADMKMLNKEKQTPLEMLK 623

RESULT 12
O88521 PRELIMINARY; PRT: 1762 AA.
AC O88521;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DE 190 KDA ANKYRIN ISOFORM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kollu A.H., Devatajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, K-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -.
DR HSSP; P55273; 1BI8.
DR InterPro; IPR000169; -.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF0791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
DR SMART; SM00005; DEATH; 1.
SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;

Query Match 25.9%; Score 301; DB 11; Length 1762;
Best Local Similarity 35.9%; Pred. No. 4.7e-18;
Matches 69; Conservative 36; Mismatches 87; Indels 0; Gaps 0;

QY 20 EELKESILADKSLATRTDQDSRTALHWACSGHTEIVEIFLQLGVPVNDKDDAGWSPLHI 79
Db 539 EDVAAFLLDHGAFSLITTKGFTPLHVAAYKGLKLEVASLLQKSPDAAGKSGLTPLHV 598
QY 80 AASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDADKHYE 139
Db 599 AAHYDNQKVALLLDQOGASPHAAKNGYTPPLHAAKKNQMDIATSLLEYGADANVTRQG 658
QY 140 ATAMHRAAAKGNLKMTHILYYKASTNIQDTEGTPHLACDEERVEEAKLLYSOGASIY 199
Db 659 IASVHLAAQEGHVDMSLLSRNANVLSNKSGLTPLHLGQEDRVNVAEVLNQGAHVD 718
QY 200 IENKEEKTPLQV 211
```

```
Db 719 AOTRMGYTPLHV 730
RESULT 13
Q9NC8P8 PRELIMINARY; PRT; 1159 AA.
AC Q9NC8P8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The 11-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL: AF190635; AAF73309.1; -
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000906; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

Query Match 25.5%; Score 297; DB 5; Length 1159;
Best Local Similarity 29.7%; Pred. No. 6.le-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 25 SILADK--SLATRTDQDSRTALHWACSAAGTTEIVEFLQLQGVVNDKDGWSP----- 76
Db 255 SLLLEKGGNIEAKT-RDGLTFLHCAARSGHEQVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 77 -----LHIAASAGRDEIVKALLGKGAQVNAVNGGCTP 109
Db 314 GEHVDAARILLYHRAPVDEVTVDTLALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 110 LHAAASKNRHEIAVMLLEGGANPDADKHYEATAMHRAAAKGNLMIHILLYKASTNIQD 169
Db 374 LHIACKNRKLVKVELLRLHSGASISATTESGLUTPLHVAAFMCGMNVIVYLQHDASPDVPT 433
Qy 170 TEGNTPPLHLACDEERVEREAKLLVSGASIIYIENKEETPLOVAK--GGIGLILKRMVSG 226
Db 434 VRGETPLHLARANQTDIIRILLRNGAODARAREEQPTPLHIASRLGNVDIVMLLLQHG 492

RESULT 14
Q9VSA2 PRELIMINARY; PRT; 2443 AA.
AC Q9VSA2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ESTS:149B10S PROTEIN.
GN ANK2 OR ESTS:149B10S OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003558; AAF50525.1; -.
DR HSSP; P42773; 11HB.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 2443 AA; 269538 MW; 7561DC5CF56812DC CRC64;
```

Query Match 25.5%; Score 297; DB 5; Length 2443;
Best Local Similarity 29.7%; Pred. No. 1.7e-17;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

```
Qy 25 SILADK--SLATRTDQDSRTALHWACSAAGTTEIVEFLQLQGVVNDKDGWSP----- 76
Db 255 SLLLEKGGNIEAKT-RDGLTFLHCAARSGHEQVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 77 -----LHIAASAGRDEIVKALLGKGAQVNAVNGGCTP 109
Db 314 GEHVDAARILLYHRAPVDEVTVDTLALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 110 LHAAASKNRHEIAVMLLEGGANPDADKHYEATAMHRAAAKGNLMIHILLYKASTNIQD 169
Db 374 LHIACKNRKLVKVELLRLHSGASISATTESGLUTPLHVAAFMCGMNVIVYLQHDASPDVPT 433
```

QY 170 TEGTPLHLACDEERVEAEKLLVSGQASIIYENKEKTPLOVAK--GGLGLILKRMVEG 226
| ||||| : : : : : || : : ||||| : | : : : : |
Db 434 VRGETPLHLAARANQTDIIIRILLRNGAQVDARAREQOTPLHLIASRLGNVDIVMLLLQHG 492

RESULT 15

Q9N180
ID Q9N180 PRELIMINARY; PRT; 1136 AA.
AC Q9N180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 1 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRED JAPANESE BLACK; TISSUE=BONE MARROW;
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT "Cloning of bovine erythrocyte ankyrin";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF222766; AAF61702.1; -
DR InterPro; IPR000906; -
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REGION; 1.
DR SMART; SM00248; ANK; 1.
FT NON_TER 1136 1136
SQ SEQUENCE 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;

Query Match 25.0%; Score 290.5; DB 6; Length 1136;
Best Local Similarity 37.5%; Pred. No. 2.3e-17;
Matches 72; Conservative 30; Mismatches 81; Indels 9; Gaps 2;

QY 26 ILADKSLATRTDODSR-----TALHWACSAGHTEIYEFLLQLGVNPNDRKDDAGWSPLHIA 80
Db 355 VLLDKG-----AKPNRSLNGFTPLHIACKKNHIRMVLELLKMGASIDAVTESGLTPLHVA 410
QY 81 ASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDAKDHYEA 140
Db 411 SFNGHPPIVKSLLQREASPNVSNKVETPLHMAARAGHTEVAKYLLQNKAQVNAKAKDDQ 470
QY 141 TAMHRAAKGNLKMIIHLLYYKASTNIQDTEGNTPLHLACDEERVEAEKLLVSGQASIIYI 200
Db 471 TPLHCAARIGHTNNVKKLLLENNANPNLATAGHTPLHIAAREGHVETALALLEKEASQTC 530
QY 201 ENKEEKTPLQVA 212
Db 531 MTRKGFTPLHVA 542

Search completed: August 13, 2001, 07:47:24
Job time: 180 sec